A CONSCIENTIOUS RESOLUTION OF THE ACTION PARADOX ON BURIDAN'S BRIDGE*

Joseph W. Ulatowski

Let us suppose that a bridge spans the width of a river connecting two separate lands. Plato and a few of his cronies stand guard at one end of the bridge. No one can cross the bridge without his assent. Socrates arrives at the bridge and pleads with great supplication for Plato to let him cross. Plato flies into a rage and swears an oath in these terms: "Surely, Socrates, if in the first proposition which you utter, you speak the truth, I will permit you to cross. But surely, if you speak falsely, I shall throw you in the water." Socrates replies, "You will throw me into the water."[1]

The dilemma is that if Plato throws Socrates in the water, then the first proposition Socrates uttered was true, in which case, to keep his promise, Plato should instead have let him freely pass. If Plato allows Socrates to cross, then the first proposition Socrates uttered was false, so that in accord with his vow, Plato should instead have thrown him in the water. This is the paradox of Buridan's Bridge. The paradox raises the question: "What ought Plato do to keep his promise?"[2]

The aim of this paper is to offer a critical assessment of Buridan's proposed solution to the bridge-keeper paradox. First, I will outline his proposed solution to the paradox, and, second, carefully analyze each issue mentioned in the proposed solution. Finally, I will attempt to conclude that Buridan has implicitly accepted a three-valued logic that does not allow him to conclude that Plato ought not do anything.

The Paradox of Buridan's Bridge and Buridan's Proposed Solution

Buridan raises three questions to provide a solution to the paradox. The first question asks for the truth-value of Socrates' proposition; likewise, the second asks whether Plato's conditional promise is true or false. The last question is the fundamental question of the paradox: "What ought Plato do in order to fulfill his promise?"

Buridan concedes Socrates' proposition is neither determinately true nor determinately false because its truth-value depends on whether a future state-of-affairs obtains. He terms this a "future contingency," which is neither

true nor false since the actions that deem the proposition true or false have not yet occurred. For this reason, Buridan believes it is in Plato's power to determine the truth-value of Socrates' proposition. The proposition is true if Plato throws Socrates in the water and false if he lets Socrates cross.

Plato's proposition is a conditional promise. When a conditional proposition conveys a promise, the antecedent asserts that something is or will be the case, and the consequent asserts that the speaker will fulfill the promise by performing some action. Buridan distinguishes between two types of conditionals—the *strict* sense and the *less strict* sense.

Strict conditionals require that the consequent be logically deducible from the antecedent condition. The consequent of Plato's conditional promise does not necessarily follow from the antecedent condition, so the antecedent could be true without the consequent being true. According to the *strict* sense, Plato's conditional promise must count as false. On the other hand, conditionals in the *less strict* sense determine the truth-value based upon the speaker either performing or not performing the action specified in the consequent. Plato by not doing anything renders the conditional false.

Now that Buridan has answered the first two questions, he turns to the fundamental question of the paradox: "What ought Plato do to fulfill his promise or vow?" Buridan appeals to the principle that what one cannot do, one has no obligation to do, which implies by contraposition the Kantian principle, "ought implies can."[3] According to Buridan, Plato cannot do anything to fulfill his promise without violating it in some way, so he ought not to keep his promise at all. He adds that Plato should not have made the promise in the first place except with some suitable qualification that would have insured the possibility of keeping the promise; the exception being that if Socrates utters some proposition that reflects upon the promise, then Socrates' proposition makes the promise impossible to fulfill.

Critical Evaluation of Buridan's Bridge

Two fundamental issues pervade the paradox of Buridan's bridge. First, we must determine the truth-value of Socrates' proposition and the truth-value of Plato's conditional promise. Let us call these questions the "logico-semantic" issues. Second, we must determine whether Plato is morally obligated to either throw Socrates in the water or allow him to pass. Let us call this the "moral" issue.

a. The Logico-semantic Issues

The logico-semantic issues are the focus of Buridan's analysis and the

Joseph W. Ulatowski

core of the paradox. He concludes (a) that Socrates' proposition is not determinately true or false, (b) neither consequent of Plato's conditional promises logically follow from the antecedent, and (c) based upon Plato not taking any action, neither Socrates nor Plato's propositions are true.

Socrates' proposition, "you will throw me into the water," refers to some state-of-affairs that may or may not occur some time in the future. For this reason, Buridan believes the statement cannot be determinately true or false. He subsequently abandons determining the truth-value of Socrates' proposition by saying, "it is in Plato's power [to determine] whether [Socrates' proposition] is true or false."[4]

That only Plato has the authority to determine the truth-value of Socrates' proposition seems counter-intuitive if for no other reason than why should Plato determine the truth-value of Socrates' proposition. A reason why Buridan may have resorted to such an extreme is his inadvertent acceptance of the Aristotelian view of truth which asserts that a sentence is true if and only if the way things are now verifies it. Such an account of truth accepts the principle of bivalence,[5] but one problem associated with future contingencies is reconciling the principle of bivalence with the view that some claims about the future are contingent. If some claims about the future are contingent, then neither the claim nor its denial is necessarily true, thereby removing the principle of bivalence.

When Buridan says that Socrates' proposition is a "future contingent" and that it is "neither determinately true nor determinately false," he implicitly denies the principle of bivalence. In doing so, he seems to accept a three-valued logic where propositions are true, false, or indeterminate. When a proposition is indeterminate, Buridan believes that the truth-value of the proposition," you will throw me into the water" cannot be verified.

Three-valued logic accepts the conclusion that some statements can be neither accepted nor rejected at the present time. This is different from saying that some propositions can be neither verified nor falsified at the present time.[6] The truth-value of a proposition cannot be identified with verification since a proposition may be verified at one time and not at another. If a proposition is ever accepted as verified, then it must be said to have been true also at times when it was not verified.[7] If Buridan accepts three-valued logic, then he must nevertheless assign an indeterminate truth-value to Socrates' proposition.

Let us suppose that Socrates' proposition has an indeterminate truth-value; what about the truth-value of Plato's conditional promise? Plato's conditional promise requires him to either allow Socrates to cross the bridge

if the first statement he utters is true or forcibly eject Socrates into the water if the first statement he utters is false. Buridan believes that Plato's conditional promise is false for two reasons: the consequent does not logically follow from the antecedent, and Plato performs neither of the actions prescribed by the conditional promise.

A strict interpretation of a material conditional proposition, such as if p then q, says that the conditional is true only if q logically follows from p. In this manner, both of Plato's conditionals must count as false: that "Plato will allow Socrates to cross" or "Plato will throw Socrates in the water" is not logically deducible from Socrates' next utterance being either true or false. The conditional cannot count as false just in case q does not logically follow from p. A conditional is false *only if* the antecedent is true and the consequent false. When the antecedent of a conditional is false, the conditional is trivially true.

For Buridan to jump to the conclusion that Plato's conditional in the strict sense is false may be a careless mistake on his part. He claims Plato's conditional promise *could not be true*, and his justification is that the antecedent *could* be true and the consequent false. That the antecedent is true and the consequent false does deductively invalidate the argument. Although this shows that the conditional promise is not necessarily true, it does not establish Buridan's stronger claim that the proposition is *necessarily not true*.

Buridan's analysis of less strict conditionals only adds confusion to settling the paradox. "A promissive conditional in a less proper sense," he asserts, "is conceded in the case that, the condition being fulfilled, the promise is fulfilled."[8] When the condition is satisfied and the promise fulfilled, the speaker who makes the conditional promise can be said to have spoken truly. Socrates does not satisfy the antecedent condition because he uttered a proposition that is neither determinately true nor determinately false. This renders the antecedent of Plato's oath unfulfilled and the entire conditional false. That the antecedent condition of a conditional promise remains unfulfilled does not necessarily falsify the promise. These rules are incomplete in that we are given no rules for the truth or falsity of the conditional when the antecedent is not true.

b. The Moral Problem

The logico-semantic problems are the crux of Buridan's bridge paradox. Buridan's solution that Plato "ought not to keep the promise" reveals the paradox's moral problem. Buridan concludes that (a) Plato is under no obligation to fulfill his vow because he carelessly formulated it without taking into account an exception and (b) Plato has simply made a false promise.

Joseph W. Ulatowski

Plato, according to Buridan, has no obligation to fulfill his vow because he has carelessly formulated it. Promises rely upon the good faith of both the promisor and the promisee to fulfill the conditions of a promise. If promises are made in good faith but may be rescinded whenever exceptions are made *ex post facto*, the institution of promising breaks down. Suppose for the moment that Buridan is right that Plato has no obligation to fulfill his vow because it is carelessly formulated. The paradox is driven by the fact that Plato did make the promise rather than whether he ought to have made the promise in the first place. It is insupportable to suppose that persons are obligated *only* to fulfill promises they ought to have made. That Plato should have instead made a promise with necessary exceptions provides a lame excuse for him not to live up to his word.

Buridan also seems to imply that Plato has made a false promise. A false promise usually means a promise that the speaker does not intend to keep. If, for instance, Wimpy says to Popeye, "I will pay you Tuesday for a hamburger today" with no apparent intention of paying Popeye for the hamburger, then Wimpy has made a false promise. This is far different than saying that a conditional statement, and in particular a conditional promise, is false. A "false promise" refers to the speaker's intentions whereas "a conditional statement is false" asserts a truth-value for a proposition.

Toward a Resolution of Buridan's Bridge

Both of Buridan's suggested solutions to the bridge paradox are inadequate. The *logico*-semantic issue jumbles several things together and the moral issue refutes exactly what drives the paradox.

Some may try to solve the paradox of Buridan's bridge by invoking either Russell and Whitehead's "ramified theory of types" or Tarski's "semantic hierarchy of truth."[9] Current orthodoxy contends that these two theories seemingly offer an adequate solution to the liar paradox; yet the two theories cannot resolve Buridan's Bridge paradox. The aspect that makes Buridan's Bridge different from the liar paradox is the offending construction of Socrates' proposition.

Socrates' reply to Plato's conditional promise does not violate either the theory of types or semantic hierarchy restrictions. Socrates' proposition is too simple syntactically to involve the application of a term of a certain

order to another term of the same order, as in the self-non-application of a predicate to that same predicate in Russell's set-theoretical paradox and its propositional counterparts.[10] The proposition also excludes it from the jurisdiction of ramified type theory since it includes no impredicative definition of terms that violate the vicious circle principle.

Moreover, Tarski's semantic hierarchy theory prohibits propositions from predicating their own truth-values. Socrates' utterance does not contain any semantic terms but only predicts what Plato's actions will be. Plato's conditional promise does indeed contain the terms "true" and "false," but Plato's conditional promise refers to Socrates' first proposition, not to itself. Semantic and moral difficulties ultimately arise only because Socrates' first proposition predicts what Plato's actions will be according to whether Socrates' proposition is true or false; however, neither Plato's conditional promise nor Socrates' proposition taken separately or in conjunction violate Tarskian restrictions on truth predication within a language.

If ramified type theory and semantic hierarchy restrictions cannot solve the paradox, then perhaps the solution proposed by Dale Jacquette may solve it.[11] Jacquette offers the conclusion that "Plato can either permit Socrates to pass or have him seized and thrown into the river *without violating his conditional vow*."[12] Plato said what he would do only in the case that Socrates' next utterance were true *simpliciter* or false *simpliciter*, but has left open what he would do if Socrates' first utterance were not true *simpliciter* or false *simpliciter*. Jacquette raises two parallel examples for comparison. He asks what would be the case if Socrates "had remained as silent as the Sphinx" or uttered an "unproven and undisproven proposition . . . [such as] Goldbach's conjecture."[13] Since Socrates can utter a proposition that is neither true *simpliciter* nor false *simpliciter*, the question of Plato acting in accord with his conditional promise does not arise. Jacquette believes that Socrates' proposition falls outside the parameters of Plato's conditional promise and does not satisfy either of the two conditions specified. This way Plato's promise is preserved as true, as something he can consistently fulfill, without reverting to the Buridanian assessment that Socrates fails to fulfill either condition of Plato's two-part conditional promise.

Jacquette's argument is at least sound from the standpoint of modern logic but wholly inadequate with respect to the paradox itself. He still does not offer an answer to the fundamental question, "What ought Plato do to fulfill his promise?" Instead, Jacquette appears to wave a wand making magical inductions about how to absolve Plato's conditional promise from Socrates' seemingly unsound proposition. If Jacquette is correct,

Joseph W. Ulatowski

then Plato can do anything he wishes to do without worrying whether he violates his own promise. His solution concludes that Plato's proposition is true, and Socrates' proposition is neither true *simpliciter* nor false *simpliciter*.

Traversing Buridan's Bridge

The problem of Buridan's bridge is easily definable, but what is not apparent is which premise must be disallowed and why it must be disallowed in order to answer the fundamental question of the paradox. Buridan does not clarify which premise he rejects. It appears that Buridan rejects both Socrates' proposition and Plato's conditional promises, yet he does not make it clear whether this is so.

Let us suppose in accord with Jacquette's assessment that Plato's conditional promise is true but that Socrates' proposition is neither true *simpliciter* nor false *simpliciter*. This seems like a plausible starting point;[14] however, we must indicate the truth-value of Socrates' proposition in order to solve the paradox.

Buridan implicitly accepts a three-valued logic when he claims Socrates' proposition is a "future contingent" that is neither determinately true nor determinately false. The reader is left with the question, "What is the truth-value of Socrates' proposition?" Buridan's acceptance of three-valued logic leads to the conclusion that Socrates' proposition has an indeterminate truth-value. Thus, there seems to be no viable option for Plato since he has not anticipated Socrates' prediction.

Buridan rejects the claim that Socrates' proposition is either determinately true or determinately false because it refers directly to a future state-of-affairs which renders the truth-value of Socrates' proposition "indeterminate." Lukasiewicz has argued something similar:

> I can assume without contradiction that my presence in Warsaw at a certain moment of next year, e.g., at noon on December 21, is at the present time determined neither positively nor negatively. Hence, it is possible, but not necessary, that I shall be present in Warsaw at the given time . . . . On this assumption, the proposition "I shall be in Warsaw at noon on December 21 of next year" can at the present time be neither true nor false.[15]

That a future state-of-affairs does not obtain is no reason to say that a proposition is not determinately true or determinately false; however, we may conclude that a proposition is indeterminate if the actions that

would satisfy the proposition have not yet occurred. If the truth-value of a proposition is indeterminate, especially in the case of Socrates, then the truth-value of Plato's conditional promise and Socrates' proposition are indeterminate. Plato ought to err on the side of caution with respect to the future contingency and allow Socrates to cross the bridge.

Perhaps a romantic conclusion to Buridan's bridge paradox has Plato and Socrates forcibly ejecting Buridan from the bridge for an eternity in the hereafter. An equally quixotic conclusion has Plato allowing Socrates to cross the bridge, but Plato pursues Socrates across the bridge and eventually tosses him into the water at the other end. A scholar of Plato may even debate with Buridan whether he means the "historical" Socrates or the "Platonic" Socrates. To my mind, the only plausible conclusion is that if Buridan accepts a three-valued logic, as he apparently does. Plato should then be merciful toward Socrates allowing him to cross sparing him from the abysmal descent into the frigid waters below since the truth-value of Plato's conditional promise and Socrates' proposition seems to be indeterminate.

NOTES

* I would like to thank Bob Barnard, Kamper Floyd, Allan Hillman, Bill Lawhead, Michael Potts, Robert Westmoreland, Scott Wilson, and members of the audience at the 2002 New Mexico-West Texas Philosophical Society Meeting for their helpful comments. Michael Harrington and Glenn Joy deserve a special note of gratitude for their careful evaluation of a draft of this paper.

[1]Jean Buridan, "Sophismata, Chapter VIII: Insolubles," in Andrew B. Schoedinger, ed., *Readings in Medieval Philosophy* (New York: Oxford UP, 1996) 731.

[2] Versions akin to this paradox have appeared elsewhere in literature and philosophy but none so clearly articulated as in Buridan's example. For example in Cervantes's *Don Quixote,* Sancho Panza must uphold a curious law that stipulates visitors upon arrival at the bridge are to be questioned about their plans and to be hanged if they answer falsely. One visitor responds, "I am here to be hanged." The dilemma is: if Sancho Panza hangs the visitor, then the visitor has answered truly and should not be hanged; moreover, if Sancho Panza does not hang the visitor, then the visitor has answered falsely and should be hanged. The similarities between this paradox and Buridan's bridge keeper paradox are incorrigible. For further information pertaining directly to Cervantes's version of the bridge keeper paradox, see Nicholas Rescher, *Paradoxes: Their Roots, Range, and Resolution* (Chicago, Open Court, 2001) and Stephen Read, "Well, I'll be Hanged! The Semantic Paradoxes," *Thinking About Logic* (Oxford: Oxford UP, 1994) 148-172.

[3] G.E. Hughes provides an extended commentary regarding Buridan's contrapositive usage of the Kantian principle. See G.E. Hughes, *John Buridan on Self-Reference: Chapter Eight of Buridan's Sophismata, with a Translation, an Introduction, and a Philosophical Commentary* (Cambridge: Cambridge UP, 1982) 222-223.

[4]Buridan 731.

[5]The principle that for every sentence *p,* either *p* is true or *p* is false.

[6] See Hilary Putnam, "Three-valued Logic," *Mathematics, Matter and Method: Philosophical Papers,* vol. 1 (Cambridge: Cambridge UP, 1975) 167.

[7]Putnam uses the example of "Christopher Columbus crosses the ocean blue in 1492." According to Putnam, this statement *was* true in 1300 A.D. and *will be* true in 5000 A.D. despite the fact that it was verified in 1492 A.D.

[8] Buridan 731.

[9] Alfred North Whitehead and Bertrand Russell, *Principia Mathematica*, rev. ed. (Cambridge, UK: Cambridge UP, 1935), vol.1, 2, B 386-417. See also Bertrand Russell, "Mathematical Logic as a Theory of Types," in Russell, *Logic and Knowledge* (London: Routledge Press, 1995) 50-73. Alfred Tarski, "The Concept of Truth in Formalized Languages," *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*, trans. J.H. Woodger (Oxford: Clarendon Press, 1956) 152-278.

[10] For a discussion of set-theoretic paradoxes and their relation to semantic paradoxes, see Alejandro R. Garciadiego, "The 'Semantic Paradoxes'" in Garciadiego, *Bertrand Russell and the Origins of the Set-Theoretic 'Paradoxes'* (Basel, Switzerland: Birkhäuser Verlag, 1992) 131-150.

[11] Dale Jacquette, "Buridan's Bridge," *Philosophy* 66 (1991) 455-471.

[12] Jacquette 466.

[13] Jacquette 467.

[14] I am not trying to imply that Socrates' proposition is the exceptionable premise.

[15] See Jan Lukasiewicz, 1930, "Many-valued Systems of Propositional Logic," *Polish Logic* (Oxford,: Oxford UP, 1967) 53.