Socrates' Last Argument

MICHAEL MORRIS

The last argument of Plato’s *Phaedo* has always fascinated philosophers. Though speciously smooth, it is naggingly hard to understand, and persistent failure provokes persistent interest. But the real importance of the argument lies in what it reveals about Plato’s metaphysical theory, and about his view of the nature of metaphysical enquiry.

If we are to see exactly what is going on, we need a formal analysis of the argument, a precise location of its errors, and some explanation of how Plato came to make them. To be at all plausible an interpretation must be charitable both to Plato and to his commentators. We should show Plato’s reasoning to be as systematic as it seems to be, and make his mistakes intelligible; and we should be able to understand the repeated failure of commentators to explain what Plato is doing.

I shall argue that the argument is formally valid and displays the systematic application of fundamental principles. Its faults are not logical in a narrow sense, but rest on what seems to us a gross metaphysical confusion. Attribution of such confusion to Plato is not new; but one of the significant features of this argument is that it can be used to justify that attribution beyond reasonable doubt. This ascription of error would nevertheless seem a lapse from charity in interpretation, if we could not make the confusion seem a natural one for Plato to make. Another striking point about this passage is that it actually provides materials for an explanation of the confusion. The source of the difficulty is Plato’s mistaken semantic theory, which is all but spelled out here. On the basis of this passage, we can see that the theory of Forms is largely motivated by a concern to make sense of certain semantic assumptions, and many of its faults are traceable to faults in the semantic theory.

I divide the passage from 102a10 to 107b10 into two parts, roughly corresponding to a stylistic division in the text. The first part (102a10 - 105c8) I
call the prelude to the argument; the second part (105c9 - 107b10) I call the argument proper. The first task is to provide an analysis of the argument proper which satisfies some compromise between these three requirements: it should keep as close to the text as possible; it should represent the logical moves made as clearly as possible; and it should aim to portray the argument as valid. Problems in interpreting and evaluating the premises will then be clearly defined.

The argument proper draws on two premisses established in the prelude, which need to be included in an analysis. It can then be represented in quantifier-variable notation in a way that makes it clear that it works by very simple rules of inference. (Premisses are marked “*”. I use parentheses to help indicate scope. Italicised expressions should be read, for the time being, as singular terms. Key expressions in the predicates are capitalised, to dull the temptation to interpret them naturally.)

Here is the argument:

(1*) For any x and y, if (if x Attacks y, then y does not Receive x), then (if x Attacks y, then either y Perishes or y Withdraws) (102d8 - e2)

(2*) For any x, y, and z, if (x Brings up y and y is Opposed to z), then (if z Attacks x, then x does not Receive z) (105a3 - 5)

(3*) soul Brings up life (105d3 - 4)

(4*) life is Opposed to death (105d6 - 9)

(5) If death Attacks soul, then soul does not Receive death (105d10 - 11; from (2*), (3*), (4*) )

(6*) For any x, if (if death Attacks x, then x does not Receive death), then x is Immortal (105e2 - 3)

(7) soul is Immortal (105e6; from (5), (6*) )

(8*) For any x, if x is Immortal, then x does not Perish (106d2 - 7)

(9) If soul is Immortal, then soul does not Perish (106e1 - 2; from (8*) )

[(10) soul does not Perish (from (7), (9) )]

(11) If death Attacks soul, then soul Withdraws (106e6 - 7; from (1*), (5), (10) )