Telephone (0865) 276310
telephone (0-896)-276327
from Mr J.R. Lucas, F.B.A.

MERTON COLLEGE
OXFORD
OX1 4JD

14 February, 1989

Dear Mr Clarke,

Thank you for sending me your paper. Yes; I think it is a fresh approach, and well worth publishing, though with one or two amendments.

You need to expand the argument on p.2 a bit, or people will snipe at you for failing to observe the distinction between syntactic statements coded into Gödel numbers and semantic statements about them.

The normal formulation of the Gödelian formula, G, is not in terms of G itself, but the Gödel number of G, which I shall type with a lower case g. Standardly it is a numerical formula $-(\exists x)Pr(x,g)$, where g is the Gödel number of that particular formula, and x ranges over Gödel numbers of sequences of well-formed formulae, which I shall refer to as X, and Pr is a dyadic predicate which holds just in case that X is a proof sequence whose last well-formed formula is G. That is, Pr(x,g) is a predicate asserting a relation between numbers, while your Dem is a predicate asserting a syntactic relation between a sequence of well-formed formulae and a particular well-formed formula. Your Dem is perfectly intelligible, but some people will carp at the existential quantifier $(\exists X)$, ranging over sequences of well-formed formulae, whereas if you talk only of the coded version, in terms of Gödel numbers, it is more difficult to carp.

The next stage, the introduction of (G^G), also needs making critic-proof. You might do it in two stages: first replace the Gödel number with another variable y,

$-(\exists x)Pr(x,y)$

and then explain the y as the Gödel number of the altered Gödelian formula in which the variable x has been replaced by z: where y is the Gödel number of G (z/x). This is rather clumsy, and it might be better to keep the variable x in the final G, and have the new quantified variable, z, at the outset. In that case you would introduce (G^G) as

$-(\exists z)Pr(z,y)$, where y is the Gödel number of $-(\exists x)Pr(x,g)$, which is of course, G itself; from which it follows that $y = g$.

You then will need some argument to show that the well-formed formula $-(\exists z)Pr(z,g)$ [i.e. $-(\exists z)Pr(z,y)$ with g substituted for y] is equivalent to $-(\exists x)Pr(x,g)$, and has the same truth-value when given the intended interpretation in the natural numbers. This again is tediously obvious, but needs spelling out, because the Gödel coding is purely syntactic, and the Gödel number of two formulae differing only in a bound variable will none the less be different.

As always in dealing with Gödel's theorem, there is a problem of the wood and the trees. Your account is easier to follow, and people could well be confused by my excessive detail. You might either put some of my version as a gloss after your version (this is why I have used Pr instead of your Dem); or you might keep your version in the main text, and put an expanded version in an Appendix.

It might be worth also giving an informal argument for your (G^G): if it could be demonstrated that there is no demonstration of G, that would constitute a formal proof (in the relevant system) of 'there is no demonstration of G', i.e. of G itself; and if we could prove G, we should thereby have proved that G was unprovable, and hence that the whole system was inconsistent.

Good wishes for its success

Yours sincerely,

J.R. Lucas

J.R. Lucas

PS I glanced at my paper to see if I could find for you the exact reference to the passage quoted by Slezak in your note 3, but my eye did not light on it, and I could not spare the time to look through it properly. It may be in my The Freedom of the Will, Oxford, 1970, in which I expanded the argument, and tried to note and meet the objections I had by then come across. But time presses. J.R.L.