The Final Cut

Elia Zardini

Department of Logic and Theoretical Philosophy
Faculty of Philosophy
Complutense University of Madrid
International Laboratory for Logic, Linguistics and Formal Philosophy
School of Philosophy and Cultural Studies
Higher School of Economics

ezardini@ucm.es

July 8, 2022 / 13th Panhellenic Logic Symposium, Volos

Outline

Historical Background

- In a series of works (e.g. Cobreros et al. [2013]), P. Cobreros, P. Égré, D. Ripley and R. van Rooij (henceforth 'CERV') have proposed a nontransitive system as a basis for a solution to the semantic paradoxes.
- The system consists of a background system (henceforth '**K3LP**') which is extended to a system of truth.
- **K3LP** has fallen out as a particular case (sort of) from the general family of tolerant logics that I've introduced in Zardini [2008a]; Zardini [2008b].
- My own aim was to provide a logic suitable for naive vagueness, where failure of transitivity is utterly compelling (in fact, CERV too initially applied their system to vagueness, e.g. Cobreros et al. [2012]).
- I did briefly consider an application to the semantic paradoxes, but couldn't find a *natural reason* for thinking that transitivity fails in that context and the logics that seemed to emerge didn't have the *desired features* anyway.

Plan for This Paper

- Throughout the years, a new reason has become clear to me for why a nontransitive approach to the semantic paradoxes is problematic.
- That reason is the main topic of this paper.
- On the way to it, I'll also raise several other issues, which variously connect to the second worry I mentioned above.
- All those issues have analogues in the case of vagueness, but I'll save you the details.

The Background System

- Without going into formal details, **K3LP** can be thought of in two ways.
- *Model-theoretically*, as the relation of logical consequence arising from the strong Kleene valuation scheme by setting the standard for premises to be 1 (as in **K3**) and the standard for conclusions to be $1/2$ (as in **LP**).
- *Proof-theoretically*, by taking a *standard Gentzen-style sequent system for classical logic* and subtracting from it the metaentailment of transitivity:

  (TRANS) If $\Gamma_0 \vdash \Delta_0, \varphi$ and $\Gamma_1, \varphi \vdash \Delta_1$ hold, $\Gamma_1, \Gamma_0 \vdash \Delta_0, \Delta_1$ holds.

- Either presentation makes it clear that *the consequence relation of* **K3LP** *coincides with the one of classical logic* (since every **K3LP** countermodel can be transformed into a classical countermodel and since cut is admissible in a standard Gentzen-style sequent system for classical logic).

Logic over and below Valid Entailments

- CERV themselves like to gloss this fact by saying that **K3LP** *is* classical logic.
- By my lights, for many a philosophically interesting logic, (there is at least one theoretically fruitful understanding on which) it is *not* reducible to the set of its valid entailments.
- Going "upwards", some logic might be just as crucially characterised *by the validity of some metaentailments*: for example, (TRANS) is arguably just as defining a principle of classical logic as e.g. the entailment of reflexivity.
- Going "downwards", some logic might be just as crucially characterised *by certain properties of its operations*: for example, that the truth of a negation rules out the truth of what it denies is arguably just as essential to classical negation as the entailment of explosion.
- I'd like to exemplify the *interaction* between these two directions by briefly presenting what I regard as one theoretically fruitful understanding of classical logic.

Facts and Alternatives

- On this understanding, classical logic includes a particular conception of the logical operations: operational classicism.
- In turn, operational classicism is better understood as having a richer-than-usual conception of structural objects.
- It inherits from the mainstream the idea that sentences can be *combined* as facts or as alternatives.
- It however severs the exclusive connection (in the mainstream) of fact combination with premise combination and of alternative combination with conclusion combination, so that these two modes of combination can *freely interact* both in combining premises and in combining conclusions.

Nothingnesses and Entailment

- Operational classicism also distinguishes two senses of nothingness: *positive* nothingness and *negative* nothingness, and treats both nothingnesses as *freely combinable*.
- Operational classicism finally follows the mainstream in adopting a general format where complexes built up by the previous structural objects and sentences are always further embedded by a *nonembeddable* entailment.

Operational Classicism

- Operational classicism is then the idea that *logical operations have the function of being an object-language representation of the corresponding structural objects*.
- The function of the conjunction of φ with ψ is to represent that φ *and* ψ *are combined as facts*. Consequently, $\varphi \,\&\, \psi$ is fully intersubstitutable with φ, ψ.
- The function of the disjunction of φ with ψ is to represent that φ *and* ψ *are combined as alternatives*. Consequently, $\varphi \vee \psi$ is fully intersubstitutable with $\varphi{:}\psi$.
- The function of the negation of φ is to represent, *together with* φ *as a combined fact, negative nothingness*, and, *together with* φ *as a combined alternative, positive nothingness*. Consequently, $\varphi, \neg\varphi$ is fully intersubstitutable with $\bullet$ and $\varphi{:}\neg\varphi$ is fully intersubstitutable with $\circ$.
- The function of the implication from φ to ψ is to represent that φ *entails* ψ. Consequently, $\Gamma \vdash \Delta{:}\varphi \to \psi$ holds iff $\Gamma, \varphi \vdash \Delta{:}\psi$ does.

Structural Classicism

- Classical logic also includes a particular conception of the structural objects: structural classicism.
- As for *properties* of single structural objects, fact and alternative combinations are *associative*, *commutative* and *idempotent*; entailment is *reflexive* and *transitive*.
- As for *relations* between different structural objects, fact combination *selects* alternative combination, they *juxtapose* and are *monotonic*; positive nothingness is an *identity for fact combination* and negative nothingness is an *identity for alternative combination*.
- Restricting to arguments where premises are combined as facts and conclusions as alternatives, operational classicism *plus* structural classicism validates exactly the classically valid arguments.
- Lifting that restriction, operational classicism *plus* structural classicism validates exactly the arguments that, on any serious understanding, should be considered as "classically valid" within the richer-than-usual conception of structural objects presented here.

Deep Nonclassicality

- On the analysis of classical logic as the sum of operational classicism and structural classicism, transitivity is a defining principle of classical logic, and so, on the analysis, **K3LP** falls short of classical logic.
- Further, on the analysis, **K3LP** *falls short of classical logic not only at the structural level, but even at the operational one*.
- **K3LP** can be extended to a system where something of the form $\varnothing \vdash \varphi \,\&\, \neg\varphi$ holds, and so where, by monotonicity, for every ψ and χ, $\psi \vdash \chi{:}\varphi \,\&\, \neg\varphi$ holds.
- According to operational classicism, by the properties of conjunction, that is equivalent with $\psi \vdash \chi{:}(\varphi, \neg\varphi)$, which, by the properties of negation, is in turn equivalent with $\psi \vdash \chi{:}\bullet$.
- Therefore, **K3LP** is incompatible with the properties of conjunction or of negation appealed to above, and so falls short of classical logic not only at the structural level, but even at the operational one.
- Thus falling short on both counts, **K3LP** can actually reasonably be taken to be a deeply nonclassical system.

Superficial Nonclassicality

- While operational classicism would seem to constitute a fairly *appealing principled, orderly conception of related functions of logical operations*, structural classicism would seem to constitute a rather *unappealing unprincipled, disorderly list of sundry properties and relations of structural objects*.
- Therefore, deviations from structural classicism would seem less problematic than deviations from operational classicism, to the extent that a system deviating only in the former way can reasonably be taken to be a superficially nonclassical system.
- I myself have developed a broadly *noncontractive* system for solving the semantic paradoxes, which vindicates operational classicism and only abandons the *idempotency* component of structural classicism (Zardini [2011]; [2014b]; [2015]; [2016]; [2019a]).
- Moreover, in my works mentioned earlier (with the further developments of Zardini [2014a]; [2015]; [2019b]), I've developed a broadly *nontransitive* system for solving the paradoxes of vagueness, which vindicates operational classicism and "only" abandons the *transitivity*, *selection* and *juxtaposition* components of structural classicism.

The Truth-Theoretic System

- The interest comes when *further constraints* on admissible models are imposed (model-theoretically) or when *further principles* are added (proof-theoretically).
- In particular, *model-theoretically*, we can impose constraints on admissible models so as to have a truth predicate T such that φ *and* $T(\ulcorner\varphi\urcorner)$ *always get the same value*.
- *Proof-theoretically*, we can add inference rules that guarantee that φ *and* $T(\ulcorner\varphi\urcorner)$ *are proof-theoretically indiscriminable*.
- That yields a system (henceforth '$\mathbf{K3LP^T}$') where φ and $T(\ulcorner\varphi\urcorner)$ are fully intersubstitutable, which some may take as capturing the essence of naive truth (we'll see).
- While (TRANS) did hold in **K3LP**, it does no longer hold in $\mathbf{K3LP^T}$, so that $\varnothing \vdash_{\mathbf{K3LP^T}} \varnothing$ does not hold.

Who Cares?

- On my view, *it doesn't matter* whether $\varnothing \vdash_{\mathbf{K3LP^T}} \varnothing$ holds—have you ever heard a compelling informal presentation of the Liar paradox where *the troubling conclusion* is supposed to be that *the empty set entails the empty set* (?!), or that *everything entails everything*?
- Well, maybe something like the latter one a couple of times. But, let λ be $\neg T(\ulcorner\lambda\urcorner)$ and suppose that the only atomic sentence of the (sentential) language is $T(\ulcorner\lambda\urcorner)$.
- Then the worst thing that something can entail is presumably something along the lines of $T(\ulcorner\lambda\urcorner) \,\&\, \neg T(\ulcorner\lambda\urcorner)$.
- But that's a logical truth of $\mathbf{K3LP^T}$, so as far as $\mathbf{K3LP^T}$ is concerned there is no paradox.
- But there is—have you ever heard a compelling informal presentation of the Liar paradox where it all hinges on the fact that language can express e.g. the proposition that Aristotelian physics is true?

Elaborating the Point

- Notice that I'm not saying that a paradox must have an *apparently false* conclusion, for it must not (López de Sa and Zardini [2007], p. 246; Oms and Zardini [2019], p. 8, fn 14).
- What I'm saying (now) is that a paradox must have a conclusion *apparently unsupported by the apparently valid reasoning*.
- But that $\varnothing \vdash_{\mathbf{K3LP^T}} T(\ulcorner\lambda\urcorner) \,\&\, \neg T(\ulcorner\lambda\urcorner)$ holds is *not* such a conclusion: after all, it's got by establishing that $\varnothing \vdash_{\mathbf{K3LP^T}} T(\ulcorner\lambda\urcorner)$ and $\varnothing \vdash_{\mathbf{K3LP^T}} \neg T(\ulcorner\lambda\urcorner)$ hold—if the latter are apparently supported, so is the former.
- You can of course accept the point and say that the paradoxical conclusions are rather that $\varnothing \vdash_{\mathbf{K3LP^T}} T(\ulcorner\lambda\urcorner)$ holds and that $\varnothing \vdash_{\mathbf{K3LP^T}} \neg T(\ulcorner\lambda\urcorner)$ holds—after all, contradictions are always apparently false.
- But then the solution to the paradox is that *contradictions are sometimes true*, not that *transitivity fails*.
- That's both old news for everyone (Priest [2006]) and bad news for lovers of nontransitive classicism.

Elaborating the Point Further

- I'm happy to grant that it'd be slightly better news *if* the main reason for rejecting contradictions were that, by transitivity, they entail everything.
- But it isn't. Ask Stephen Read.
- He'll tell you that, although contradictions don't entail everything, we should reject them because *they cannot be true*.
- I'd add that they cannot be true because one (prominent) understanding of negation is such that the truth of a negation *rules out* the truth of what it denies (I'll give below a fuller statement under the label 'complementative negation').
- If you want to accept contradictions you mainly need to face this kind of challenge (as Priest [2006] admirably does), where the issue of transitivity is neither here nor there.

A Paradoxical Solution

- On my view, the fact that $\varnothing \vdash_{\mathbf{K3LP^T}} \varnothing$ does not hold also *comes too late*.
- For one thing, $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda$ and $\lambda \vdash_{\mathbf{K3LP^T}} \varnothing$ still both hold: λ is *inconsistent* but is also a *logical truth*.
- In other words, *something must be the case while at the same time it cannot be the case*.
- That in itself is something *that cannot be the case* (indeed, that's exactly the paradox!).
- Therefore, $\mathbf{K3LP^T}$ *cannot be correct* (indeed, it just reinstates the paradox!).

Condemning Convergence

- For another thing, the absolutely central feature of naive truth is the correlation principle between reality and truth, which, subject to qualifications which needn't detain us here, in its convergence-oriented version can be formulated as $\varphi \leftrightarrow T(\ulcorner\varphi\urcorner)$.
- The Liar paradox is an argument against (the convergence-oriented version of) that compelling principle, allegedly showing that it is inconsistent (or, if you really prefer, allegedly showing that it entails that Aristotelian physics is true).
- But $\lambda \leftrightarrow T(\ulcorner\lambda\urcorner) \vdash_{\mathbf{K3LP^T}} \varnothing$ holds.
- Therefore, $\mathbf{K3LP^T}$ *condemns naive truth to the Liar paradox just as the grimmest classical logician does*: for $\mathbf{K3LP^T}$ just as well, the correlation principle is inconsistent and entails that Aristotelian physics is true.

Truth-Makers for Aristotelian Physics

- It's true that $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda \leftrightarrow T(\ulcorner \lambda \urcorner)$ also holds and so, contrary to the classical logician, $\mathbf{K3LP^T}$ also maintains the correlation principle holds.
- But that only makes things worse (if possible), for it means that $\mathbf{K3LP^T}$ maintains that *something holds which entails that Aristotelian physics is true*—that *there is something that makes Aristotelian physics true*.
- With all due respect for what is in some *technical* respects an interesting system, those are ludicrous statements, which constitute abundant reason for conclusively rejecting $\mathbf{K3LP^T}$ in its envisaged *philosophical* applications.
- It's no use to say that the entailment does not really "force" the truth of Aristotelian physics, for then, by the same token, one should conclude that the entailment $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda \leftrightarrow T(\ulcorner \lambda \urcorner)$ does not really "force" that the correlation principle holds.

Condemning Nondivergence

- Similarly, in its nondivergence-oriented version, subject to qualifications which needn't detain us here, the correlation principle between truth and reality can be formulated as $\neg(\varphi \,\&\, \neg T(\ulcorner\varphi\urcorner))$ and $\neg(T(\ulcorner\varphi\urcorner) \,\&\, \neg\varphi)$.
- Of course, those are also inconsistent *etc.* in $\mathbf{K3LP^T}$ (and in virtually every *nondialetheic* approach to the paradoxes except mine, e.g. Field [2008]; see Heck [2012]; Zardini [2013a]), but I'm not going to belabour that.

Embracing Divergence

- Rather, notice that they are supposed to express the informal ideas that *it cannot be the case that what a sentence says is the case without the sentence's being true*, and that *it cannot be the case that a sentence is true without what it says being the case*.
- But $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda \,\&\, \neg T(\ulcorner \lambda \urcorner)$ and $\varnothing \vdash_{\mathbf{K3LP^T}} T(\ulcorner \lambda \urcorner) \,\&\, \neg \lambda$ hold.
- Therefore, according to $\mathbf{K3LP^T}$, it is indeed the case that what a sentence says is the case without the sentence's being true, and it is indeed the case that a sentence is true without what it says being the case.
- This is really an argument against virtually every *dialetheic* approach to the semantic paradoxes (e.g. Priest [2006]).

Paradox from 'Without'

- It is tempting to reply that $\varphi \mathbin{\&} \neg\psi$ is *not strong enough* in $\mathbf{K3LP^T}$ to mean that φ holds without ψ holding. (More because of the inadequacy of negation rather conjunction, I speculate.)
- But, if this is not meant by that, it is not meant by anything in $\mathbf{K3LP^T}$.
- And, if we try to introduce conjunction-like and negation-like operators $\&^*$ and $\neg^*$ strong enough so that $\varphi \mathbin{\&^*} \neg^*\psi$ does mean that φ holds without ψ holding, they'll have to validate the compelling "with or without you" law $\varnothing \vdash \mathsf{t} \mathbin{\&^*} \varphi \vee \mathsf{t} \mathbin{\&^*} \neg^*\varphi$ (either what is logically true holds together with φ holding or without it)...
- ...and the compelling "lonely together" metaentailment from $\varnothing \vdash \varphi$ and $\varnothing \vdash \neg^*\psi$ to $\varnothing \vdash \varphi \mathbin{\&^*} \neg^*\psi$ (if it is a logical truth that φ holds and a logical truth that $\neg^*\psi$ holds, it is a logical truth that φ holds without ψ holding).
- The above argument will then kick in (for a λ^* identical with $\neg^* T(\ulcorner\lambda^*\urcorner)$).

Without ‘Without’

- What’s the point of trying to make the notion of naive truth expressible at the cost of making the notion of withoutness—which crucially enters in the formulation of the nondivergence-oriented version of the correlation principle—*inexpressible*?
- In essence, this is a particularly strong version of the argument against dialetheism based on complementative negation (*i.e.* exhaustive and exclusive negation): the negation that holds iff what it denies *somehow or other* (exhaustion) *fails* (exclusion) to hold.

Full Intersubstitutability against Naive Truth

- Given how badly $\mathbf{K3LP^T}$ treats the correlation principle, one might wonder how virtually every well-established researcher in the area could seriously consider it among the candidate systems of naive truth.
- I'm pretty sure that's because $\mathbf{K3LP^T}$ validates the full intersubstitutability of φ with $T(\ulcorner\varphi\urcorner)$.
- But that principle is "too stronger" than naive truth and in fact variously inconsistent with it.
- For one thing, if there is *no fact of the matter* about φ, then φ does not correspond to the facts and so is not true, but $\neg\varphi$ doesn't follow.
- For another thing, another feature of naive truth is that, assuming φ, φ is true *because* φ holds. But, for the vast majority of φs (Zardini [2019a], p. 176, fn 47 provides one kind of exception), it is not the case that φ is true because φ is true.

Full Intersubstitutability for Deflationary Truth

- Given this, one might wonder how virtually every well-established researcher in the area could seriously consider full intersubstitutability a desideratum of naive truth.
- I'm pretty sure that's because of deflationist propaganda.
- But, in the only philosophically serious sense, deflationism about truth is the wildly implausible doctrine that *the truth predicate is merely an expressive device* (though there are deep philosophical arguments in its favour, see Field [1994]).
- Contrary to naive truth, deflationary truth does indeed require full intersubstitutability.

Restricted Intersubstitutability

- That said, I do think that there is a more roundabout reason for having full intersubstitutability as a desideratum of naive truth having to do with higher-order paradoxes (Zardini [2020]).
- A less roundabout reason is that *a lot of* applications of full intersubstitutability remain compelling, and we can stipulate that the language does not express any of the phenomena that make it fail, so that *all* applications of full intersubstitutability become compelling.

Restricted Intersubstitutability as Grounded in the Correlation Principle

- But those applications are compelling *in virtue of their being grounded in the correlation principle*. (Ask yourself why you believe that 'Snow is white or grass is green' entails ' 'Snow is white' is true or grass is green'.)
- The holding of full intersubstitutability in a system of naive truth should therefore be grounded in the holding of the correlation principle in the system.
- But this requirement is *not* met by $\mathbf{K3LP^T}$.
- In general, in $\mathbf{K3LP^T}$, there's precious little that is constrained about the behaviour of φ and ψ simply by the fact that $\varnothing \vdash_{\mathbf{K3LP^T}} \varphi \leftrightarrow \psi$ holds.

The Conjunction of Logical Truths

- On my view, the fact that $\varnothing \vdash \varnothing$ should hold in $\mathbf{K3LP^T}$ is also anyway *ultimately inescapable*.
- If $\varnothing \vdash_{\mathbf{K3LP^T}} \lambda$ holds, λ is a logical truth.
- Since t is the (big) *conjunction of logical truths*, it has λ among its conjuncts.
- Since even a defender of $\mathbf{K3LP^T}$ should accept that *a conjunction entails everything any of its conjuncts entails* (and since $\lambda \vdash_{\mathbf{K3LP^T}} \varnothing$ holds), $\mathrm{t} \vdash \varnothing$ should hold in $\mathbf{K3LP^T}$.
- As before, that's bad enough: it is a sort of ultrairrationalist position according to which *the laws of logic cannot be the case*.

Reaching Total Triviality

- Even worse, the deletion/insertion metaentailment that $\Gamma, t \vdash \Delta$ holds iff $\Gamma \vdash \Delta$ holds is compelling: *entailment in virtue of the laws of logic is entailment is entailment in virtue of the laws of logic*.
- Therefore, $\varnothing \vdash \varnothing$ should hold in $\mathbf{K3LP^T}$.
- A dual argument is available for logical falsehood.

Naive Validity and the Manifestation Principle

- Time to stop about truth. For many of us, another central semantic notion is validity.
- The absolutely central feature of naive validity is the manifestation principle of the *metatheoretic facts of entailment* by the *object-theoretic facts of validity*.
- With V a validity predicate, focusing on a single system and on single-premise single-conclusion validities, $\Gamma, \varphi \vdash \Delta, \psi$ holds iff $\Gamma \vdash \Delta, V(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner)$ holds, where all elements of Γ and Δ are "validity sentences".
- For the purposes of this paper, atomic V-sentences as well as t and f are validity sentences, and so are negations, conjunctions, disjunctions and implications of validity sentences (Zardini [2014b], p. 355).

A Hypothetical Connection

- Notice that, just like the correlation principle connects the facts about reality and the facts about truth not only *categorically*, but also *hypothetically*. . .
- . . . so the manifestation principle connects the facts about entailment and the facts about validity not only *categorically*, but also *hypothetically* (what entails what *under the hypothesis that all the elements of* Γ *hold and all the elements of* Δ *do not hold*).

The Naive-Validity Version of Curry's Paradox

- It is possible to extend **K3LP** to a system (henceforth '$\mathbf{K3LP^V}$') where the manifestation principle holds.
- The manifestation principle gives rise to a naive-validity version of Curry's paradox.
- Letting κ be $V(\ulcorner\kappa\urcorner, \ulcorner \mathrm{f} \urcorner)$, $\varnothing \vdash_{\mathbf{K3LP^V}} \kappa$ and $\kappa \vdash_{\mathbf{K3LP^V}} \varnothing$ hold.
- While (TRANS) did hold in **K3LP**, it does no longer hold in $\mathbf{K3LP^V}$, so that $\varnothing \vdash_{\mathbf{K3LP^V}} \varnothing$ does not hold.
- For reasons analogous to those seen in the case of λ, this is not a particularly effective solution of the paradox, but I'm not going to insist on those.
- (Shapiro [2013] touches on something like the fact that, in $\mathbf{K3LP^V}$, *something is supposed both to hold but also to entail something crazy*, but as, we've seen, that's already the case for $\mathbf{K3LP^T}$'s treatment of the Liar paradox.)

Naive Validity Falls Short of Validity in **K3LP^V**

- I'm going to develop instead a specific problem arising in connection with the naive-validity version of Curry's paradox.
- Notice that, in **K3LP^V**, both the notion of *what is valid in* **K3LP^V** and the notion of *what is naively valid* are expressible.
- Suppose now that, insofar as issues of naive validity are concerned, you accept **K3LP^V**. What relations should you think obtain between those two notions?
- In one direction, you should presumably still think that *there is a gap from naive validity to validity in* **K3LP^V**: after all, there are presumably naively valid arguments that are not valid in **K3LP^V** (for example, the argument from 'It is necessary that φ' to φ).

Validity in $\mathbf{K3LP^V}$ Does Not Fall Short of Naive Validity

- In the other direction, however, you should presumably now think that *there is no gap from validity in* $\mathbf{K3LP^V}$ *to naive validity*: given any argument, you should only suppose that it is valid in $\mathbf{K3LP^V}$ if you also suppose that it is naively valid.
- In a normal situation, the absence of such a gap would be guaranteed by an implication from an argument's being valid in $\mathbf{K3LP^V}$ to its being naively valid.
- But, as I've already noted, the situation is not normal: in $\mathbf{K3LP^V}$, implication has nothing like that force (for example, in $\mathbf{K3LP^V}$, the implication from λ to f holds, although the former is accepted and the latter rejected).

One-Way Substitutability

- The absence of such a gap can still be guaranteed by one-way substitutability: letting $V_{\mathbf{K3LP^V}}$ express validity in $\mathbf{K3LP^V}$, V can be substituted for $V_{\mathbf{K3LP^V}}$ in "upwards-monotonic contexts", and *vice versa* in "downwards-monotonic" contexts.
- For the purposes of this paper, stand-alone sentences as conclusions (premises) are in an upwards-monotonic context (downwards-monotonic context); sentences in an upwards-monotonic context (downwards-monotonic context) further embedded under negation or as antecedents of an implication are in a downwards-monotonic context (upwards-monotonic context); sentences in an upwards-monotonic context (downwards-monotonic context) further embedded as consequents of an implication, or as conjuncts of a conjunction, or as disjuncts of a disjunction are in an upwards-monotonic context (downwards-monotonic context).

A Hierarchy for One-Way Substitutability

- Now, since we may assume that $\mathbf{K3LP^V}$ contains a *classical theory of syntax*, one-way substitutability cannot be part of $\mathbf{K3LP^V}$ itself.
- The reason for this is *Gödelian*. By reflexivity and the manifestation principle, $t, V(\ulcorner t \urcorner, \ulcorner f \urcorner) \vdash_{\mathbf{K3LP^V}} f$ holds, and so, by the properties of negation in $\mathbf{K3LP^V}$, $t \vdash_{\mathbf{K3LP^V}} \neg V(\ulcorner t \urcorner, \ulcorner f \urcorner)$ holds.
- If one-way substitutability were part of $\mathbf{K3LP^V}$, it would follow that $t \vdash_{\mathbf{K3LP^V}} \neg V_{\mathbf{K3LP^V}}(\ulcorner t \urcorner, \ulcorner f \urcorner)$ holds, thereby crashing on the second incompleteness theorem.
- Therefore, we must assume that one-way substitutability is only part of a *stronger* system $\mathbf{K3LP^{VS^{\Rightarrow}}}$.

The Final Cut

- Since $V(\ulcorner\kappa\urcorner, \ulcorner f\urcorner) \vdash_{\mathbf{K3LP^V}} f$ holds, by one-way substitutability $V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner) \vdash_{\mathbf{K3LP^{VS\Rightarrow}}} f$ holds.
- Moreover, since $\kappa \vdash_{\mathbf{K3LP^V}} f$ holds and $\mathbf{K3LP^V}$ contains a classical theory of syntax, $t \vdash_{\mathbf{K3LP^V}} V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner)$ holds, and so $t \vdash_{\mathbf{K3LP^{VS\Rightarrow}}} V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner)$ holds.
- Now, presumably, even in a $\mathbf{K3LP^{VS\Rightarrow}}$-style system, (TRANS) should hold *when the cut formula is an elementary, provable mathematical fact*, such as $2 + 2 = 4$...
- ...or $V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner)$. But then we can cut on $t \vdash_{\mathbf{K3LP^{VS\Rightarrow}}} V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner)$ and $V_{\mathbf{K3LP^V}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner) \vdash_{\mathbf{K3LP^{VS\Rightarrow}}} f$, and get that $t \vdash_{\mathbf{K3LP^{VS\Rightarrow}}} f$ holds, and so, by the deletion/insertion metaentailments, that $\varnothing \vdash_{\mathbf{K3LP^{VS\Rightarrow}}} \varnothing$ holds.

The Connection between $\mathbf{K3LP}^{\mathbf{VS}^{\Rightarrow}}$ and $\mathbf{K3LP}^{\mathbf{V}}$

- Therefore, $\mathbf{K3LP}^{\mathbf{VS}^{\Rightarrow}}$ is trivial.
- At the same time, it would seem to be the system that a defender of $\mathbf{K3LP}^{\mathbf{V}}$ should adopt to make sure that, in the object theory, there is no gap from validity in $\mathbf{K3LP}^{\mathbf{V}}$ to naive validity: one can only accept $\mathbf{K3LP}^{\mathbf{V}}$ if one accepts $\mathbf{K3LP}^{\mathbf{VS}^{\Rightarrow}}$.
- Since one can't accept $\mathbf{K3LP}^{\mathbf{VS}^{\Rightarrow}}$, one can't accept $\mathbf{K3LP}^{\mathbf{V}}$ either.

What Is the Real Target of the Argument?

- Notice that, in a sense, the argument *generalises* to all systems where $\kappa \vdash \mathsf{f}$ holds.
- In turn, those are those systems where the manifestation principle holds (left-to-right) and where the metaentailment of contraction:

(CONTR) If $\Gamma, \varphi, \varphi \vdash \Delta$ holds, $\Gamma, \varphi \vdash \Delta$ holds, and, if $\Gamma \vdash \Delta, \varphi, \varphi$ holds, $\Gamma \vdash \Delta, \varphi$ holds.

is valid (in its first conjunct).

Contractive Transitive Systems

- However, the transitive systems satisfying both those conditions are shown to be straightforwardly trivial *already by the old naive-validity version of Curry's paradox*.
- For those systems, the new naive-validity version of Curry's paradox relying on one-way substitutability is an overkill.
- Indeed, it is very much arguable that, given the kind of conception of validity that—for better or worse—*contractive transitive* systems are already committed to given the *original, implication-version of Curry's paradox*, the part of the manifestation principle that they should reject is precisely the right-to-left direction (Zardini [2013b] pp. 636–638; Field [2017]).

Noncontractive Systems

- *Noncontractive* systems can afford a much tighter connection between entailment and object-theoretic notions like implication and validity, and so can validate the manifestation principle.
- On a noncontractive system $\mathbf{LW^V}$ (e.g. Zardini [2014b]), $\kappa, \kappa \vdash_{\mathbf{LW^V}} \mathsf{f}$ holds but $\kappa \vdash_{\mathbf{LW^V}} \mathsf{f}$ doesn't.
- By considerations analogous to those made for $\mathbf{K3LP^V}$, we can extend $\mathbf{LW^V}$ to a stronger system $\mathbf{LW^{VS^{\Rightarrow}}}$ where one-way substitutability holds.

The Final Contraction?

- Since $V(\ulcorner\kappa\urcorner, \ulcorner f\urcorner), V(\ulcorner\kappa\urcorner, \ulcorner f\urcorner) \vdash_{\mathbf{LW^v}} f$ holds, by one-way substitutability $V_{\mathbf{LW^v}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner), V_{\mathbf{LW^v}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner) \vdash_{\mathbf{LW^{vs\Rightarrow}}} f$ holds.
- Now, presumably, even in a $\mathbf{LW^{VS^{\Rightarrow}}}$-style system, (CONTR) should hold *when the to-be-contracted formula is an elementary, disprovable mathematical nonfact*, such as $2 + 2 = 5\ldots$
- $\ldots$or $V_{\mathbf{LW^v}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner)$. But then we can contract on $V_{\mathbf{LW^v}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner), V_{\mathbf{LW^v}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner) \vdash_{\mathbf{LW^{vs\Rightarrow}}} f$, and get that $V_{\mathbf{LW^v}}(\ulcorner\kappa\urcorner, \ulcorner f\urcorner) \vdash_{\mathbf{LW^{vs\Rightarrow}}} f$ holds.
- But, contrary to the case of $\mathbf{K3LP^{VS^{\Rightarrow}}}$, this conclusion is *wholly benign* (indeed, it's just what you'd expect from the kind of stronger system $\mathbf{LW^{VS^{\Rightarrow}}}$ is).

The Real Target of the Argument

- Therefore, *the only system that is affected by the new naive-validity version of Curry's paradox but not by the old one is a nontransitive contractive system such as* $\mathbf{K3LP^V}$.
- There is more to naive validity than the manifestation principle.
- In particular, a system should not only connect the metatheoretic facts of its entailments with the object-theoretic facts of naive validity; the driving force behind internalisation requires that the system should also (be strengthened to) connect the latter facts with the object-theoretic facts of what is valid in the system.
- Naive validity should therefore also include one-way substitutability.
- Although $\mathbf{K3LP^V}$ (sort of) solves the old naive-validity version of Curry's paradox and validates the manifestation principle, it crashes on the new naive-validity version of Curry's paradox and cannot (be strengthened to) validate one-way substitutability.
- When $\mathbf{K3LP^V}$ talks about naive validity, it does in fact track its own entailments (among others), but it can't know that.

References

Pablo Cobreros, Paul Égré, David Ripley, and Robert van Rooij. Tolerant, classical, strict. *Journal of Philosophical Logic*, 41:347–385, 2012.

Pablo Cobreros, Paul Égré, David Ripley, and Robert van Rooij. Reaching transparent truth. *Mind*, 122:841–866, 2013.

Hartry Field. Deflationist views of meaning and content. *Mind*, 103:249–285, 1994.

Hartry Field. *Saving Truth from Paradox*. Oxford University Press, Oxford, 2008.

Hartry Field. Disarming a paradox of validity. *Notre Dame Journal of Formal Logic*, 58:1–19, 2017.

Richard Heck. A liar paradox. *Thought*, 1:36–40, 2012.

Dan López de Sa and Elia Zardini. Truthmakers, knowledge and paradox. *Analysis*, 67:242–250, 2007.

Sergi Oms and Elia Zardini. An introduction to the sorites paradox. In Sergi Oms and Elia Zardini, editors, *The Sorites Paradox*, pages 3–18. Cambridge University Press, Cambridge, 2019.

Graham Priest. *In Contradiction*. Oxford University Press, Oxford, 2nd edition, 2006.

Lionel Shapiro. Validity Curry strengthened. *Thought*, 2:100–107, 2013.

Elia Zardini. A model of tolerance. *Studia logica*, 90:337–368, 2008a.

Elia Zardini. *Living on the Slippery Slope. The Nature, Sources and Logic of Vagueness*. PhD thesis, Department of Logic and Metaphysics, University of St Andrews, 2008b.

Elia Zardini. Truth without contra(di)ction. *The Review of Symbolic Logic*, 4:498–535, 2011.

Elia Zardini. It is not the case that [P and 'It is not the case that P' is true] nor is it the case that [P and 'P' is not true]. *Thought*, 1:309–319, 2013a.

Elia Zardini. Naive logical properties and structural properties. *The Journal of Philosophy*, 110:633–644, 2013b.

Elia Zardini. Evans tolerated. In Kensuke Akiba and Ali Abasnezhad, editors, *Vague Objects and Vague Identity*, pages 327–352. Springer, Dordrecht, 2014a.

Elia Zardini. Naive truth and naive logical properties. *The Review of Symbolic Logic*, 7:351–384, 2014b.

Elia Zardini. $\forall$ and ω. In Alessandro Torza, editor, *Quantifiers, Quantifiers, and Quantifiers: Themes in Logic, Metaphysics, and Language*, pages 489–526. Springer, Dordrecht, 2015.

Elia Zardini. Restriction by noncontraction. *Notre Dame Journal of Formal Logic*, 57:287–327, 2016.

Elia Zardini. Instability and contraction. *Journal of Philosophical Logic*, 48:155–188, 2019a.

Elia Zardini. Non-transitivism and the sorites paradox. In Sergi Oms and Elia Zardini, editors, *The Sorites Paradox*, pages 168–186. Cambridge University Press, Cambridge, 2019b.

Elia Zardini. Neoclassical logic, cloisonnist vagueness and nonobjectual truth. ms, 2020.