Fitch's Paradox Beyond Verificationism

A Puzzle for Closure and Some Modally Qualified Principles

Sebastian Liu

1. Fitch's Paradox

According to strong verificationism, all truths are known.

(SVER) $p \to Kp$

Read 'Kp' as 'you know that p'.

According to weak verificationism, all truths are possibly known.

(WVER) $p \to \Diamond Kp$

Read '$\Diamond p$' as 'p is logically possible'.

Both SVER and WVER are false. SVER is false because there are true but unknown propositions. And WVER is false because, given that there are true but unknown propositions, WVER entails SVER.

For example, that I have exactly $100,000$ hairs on my head (supposing it is true) is presumably unknown. That I have exactly $100,000$ hairs on my head and I don't know it is unknowable.

Proof: Suppose WVER. Let p be true but unknown. Then $p \wedge \neg Kp$ entails $\Diamond K(p \wedge \neg Kp)$. But suppose $K(p \wedge \neg Kp)$. Since knowledge distributes over conjunction, knowing a conjunction entails knowing each conjunct, so $Kp \wedge K\neg Kp$. But knowledge is factive. So $Kp \wedge \neg Kp$. Contradiction. Therefore $\neg \Diamond K(p \wedge \neg Kp)$, and so $\neg(p \wedge \neg Kp)$, or equivalently, $p \to Kp$.

More precisely: Distribution is the assumption that $K(p \wedge q) \to (Kp \wedge Kq)$ and factivity is the assumption that $Kp \to p$.

Fitch's Paradox: WVER collapses into SVER. This is surprising!

See Fitch 1963. For details (and a strengthening of the result) see Facts 1 and 2 in the Appendix.

Observation: Nothing in the formal result depends on verificationist ideas. What matters is only that the modal operator (in the verificationism case, the knowledge operator) has certain properties.

More on these properties later.

Present aim: Show that Fitch's Paradox has implications for a wide-range of principles. I'll be focusing on the closure principle.

Hart (1979, p.164) calls Fitch's Paradox an "unjustly neglected gem".

2. Strong and Weak Closure

According to strong closure, you know everything that follows from your knowledge:

(SCLO) If you know p, and p entails q, then you know q.

Formally, this can be captured by the rule of inference RM: $p \to q / Kp \to Kq$. See e.g. Hintikka 1970 and Schaffer 2007.

SCLO appears very strong. One consequence of it is that if you know anything at all, then you know all theorems of logic and mathematics.

According to weak closure, it's possible to know everything that follows from your knowledge:

(WCLO) If you know p, and p entails q, then it is possible for you to know q.

Call this RM$^{\Diamond}$: $p \to q / Kp \to \Diamond Kq$.

WCLO appears very weak. It doesn't have the consequence that you know all theorems of logic and mathematics.

A Fitch-style argument: If you think that SCLO is false, then you think that there's some p and q such that p entails q, you know p, but you fail to know q. Presumably you can know of some p and q, that $p \wedge \neg Kq$. But $p \wedge \neg Kq$ entails $q \wedge \neg Kq$. So by WCLO, $\Diamond K(q \wedge \neg Kq)$. Contradiction.

See Fact 3 in the Appendix. See also Salow manuscript.

Upshot: WCLO entails that it's impossible for you to know $p \wedge \neg Kq$ whenever p entails q. But:

That is, given WCLO, $\neg\Diamond K(p \wedge \neg Kq)$, whenever p entails q. This can be understood as the claim that, necessarily, given that you have any knowledge, you can't know that you're not omniscient.

1. I know $p \vee \neg p$, from which the truth of the twin prime conjuncture follows. Presumably I can also know that I don't know whether the twin prime conjecture is true.
2. From any p, it follows that $p \vee q \vee r \vee \ldots$ But you can know that you don't know this disjunction by knowing that you don't even grasp it.
3. Lois Lane knows that Superman can fly, from which it follows that Clark Kent can fly. Yet she's arguably able to know that she doesn't know that Clark Kent can fly.

So: To the extent that it's possible to know, of *some* p and q (where p entails q) that $p \wedge \neg Kq$, WCLO should be rejected.

Note: The claim *isn't* that WCLO collapses into SCLO.

3. *An Even Weaker Closure Principle?*

Why should we go for WCLO? The most familiar kind of closure principle states that you know everything that you *know* follows from your knowledge:

> (CLO) If you know p and you know that p entails q, then you know q.

We can generalize this. Consider:

> ... if S knows p, competently deduces q, and thereby comes to believe q, while retaining knowledge of p throughout, then S knows q. (Hawthorne 2004, p.34)
>
> ... if S knows that P is true and knows that P implies Q, then, evidentially speaking, this is enough for S to know that Q is true. (Dretske 2005, p.27)
>
> For any agent who is able to apply the relevant deductive rules to the premise that p, if she knows that p, and q is an obvious logical consequence of p, then she is in a position to know that q. (Das and Salow 2018, p.7)

There are broadly two strategies: strengthen the antecedent or weaken the consequent. Here's a generalized closure schema that captures all of these qualifications:

(GCLO) If you know p, and p entails q, and the set of conditions in X obtain, then it is possible for you to know q.

Sometimes the set of conditions in X is explicitly left vague. See e.g. David and Warfield 2008 and Immerman 2020. The set of conditions is meant to be supplied by specific versions of closure.

WCLO modifies SCLO by weakening the consequent whereas GCLO modifies SCLO both by weakening the consequent and strengthening the antecedent.

Question: Does GCLO collapse into WCLO?

Answer: No. Let p be the proposition that no one will deduce anything and let q be the proposition that I will not deduce anything. Then while it's possible for me to know p and know q, I couldn't come to know q on the basis of *deducing* it from p, for necessarily, if I were to make this deduction, I wouldn't know that no one will deduce anything. But the pair of propositions $p \wedge \neg Kq$ and $q \wedge \neg Kq$ makes WCLO false and GCLO true, so the latter doesn't always collapse into the former.

Therefore WCLO isn't the *weakest* form of closure. I only call it 'weak closure' to draw the parallel between closure principles and verificationist principles.

To be explicit: This pair of propositions makes GCLO true by making its antecedent (necessarily) false.

Say that a proposition p is *X-pathological* if and only if knowing p is incompatible with satisfying the conditions in X.

Formally: p is X-pathological iff $\neg\Diamond(Kp \wedge X)$.

Claim: Whenever $\Diamond(Kp \wedge X)$, GCLO entails WCLO. GCLO does not collapse into WCLO only when p is X-pathological.

See Fact 4 in the Appendix.

Question: What's the X needed to make all Fitch-propositions of the form $p \wedge \neg Kq$ X-pathological?

Answer: A condition that explicitly restricts against all and only Fitch-propositions?

Compare: $p \rightarrow \Diamond Kp$, provided p isn't a Fitch-proposition. Though see e.g. Dummett 2001 and Douven 2005.

Pessimism: Suppose you think that $K(p \wedge \neg Kq)$ is incompatible with 'competently deducing' $q \wedge \neg Kq$ from $p \wedge \neg Kq$. The proposal is that necessarily, if you competently deduce $q \wedge \neg Kq$ from $p \wedge \neg Kq$, then $\neg K(p \wedge \neg Kq)$. But why think this? Here's what appears to be the most plausible suggestion:

I'll abbreviate 'competent deduction of q from p' as $D(q, p)$.

($*$) If Kp and $D(q \wedge \neg Kq, p \wedge \neg Kq)$ then $\neg K\neg Kq$.

Since $K(p \wedge \neg Kq)$ entails Kp and $K\neg Kq$, if knowing p and making the relevant deduction guarantees $\neg K\neg Kq$, then you couldn't deduce $q \wedge \neg Kq$ from $p \wedge \neg Kq$ while knowing $p \wedge \neg Kq$.

Unfortunately ($*$) is problematic. For either $\neg K\neg Kq$ entails $\Diamond Kq$ or it doesn't. But:

1. If $\neg K\neg Kq$ doesn't entail $\Diamond Kq$, then GCLO fails for unrelated reasons.

   Why? If $\neg K\neg Kq$ but $\neg\Diamond Kq$, then merely restricting against Fitch-propositions won't be enough to salvage GCLO, since it's assumed that Kp and the conditions in X are satisfied, but yet $\neg\Diamond Kq$.

Note that q itself isn't assumed to be a Fitch-proposition.

2. If $\neg K\neg Kq$ does entail $\Diamond Kq$, then $(*)$ violates a putative circularity constraint, namely, that the set of conditions in X shouldn't itself rely on an instance of GCLO to guarantee that all Fitch-propositions are X-pathological.

   Why? If $\Diamond Kq$ whenever $\neg K\neg Kq$, then $(*)$ entails that whenever you know p and competently deduce $q \wedge \neg Kq$ from $p \wedge \neg Kq$, then $\Diamond Kq$. But this is an instance of GCLO, so it fails the non-circularity constraint.

Presumably if you're interested in whether GCLO is true, it would be inappropriate for you to rely on an instance of the principle to ban all Fitch-propositions that would be problematic for GCLO.

This example assumed that the condition is competent deduction, but nothing essentially depended on that assumption. So: Either there's reason to be pessimistic about a principled condition guaranteed to be incompatible with $K(p \wedge \neg Kq)$, or GCLO fails for other reasons.

We can replace $D(q \wedge \neg Kq, p \wedge \neg Kq)$ with some generic condition x for a generalized version of the argument.

Challenge: Is there a non-circular, independently motivated condition that makes all Fitch-propositions pathological?

Maybe: It also shouldn't vindicate verificationism!

4. On Other Modally Qualified Principles

This style of Fitch-argument generalizes. Most straightforwardly:

(KK) If you know p, then you know that you know p.

(EQUIV) If you know p, and p and q are logically equivalent, then you know q.

I'll be omitting the details of these proofs because they're similar in style to the ones in the Appendix. See also Liu 2020 and San 2020.

These principles, in their strongest forms, look implausible.

For instance, if knowledge requires belief, couldn't you know p without knowing that you know p by failing to have beliefs about your knowledge? See e.g. McHugh 2010 and Greco 2014 for qualified version of KK.

When the formal language is suitably amended, the argument bears on principles including:

(TESTIMONY) If I know p, and I tell you that p, then you know p.

(RECALL) If you know p at time t, then you know p at time t', where $t < t'$.

In particular, we'll need to index knowledge to particular agents and times. On these principles, see especially Fraser and Hawthorne 2015.

Fitch's Paradox relies on the modal operator O satisfying:

(DISTRIBUTION) $O(p \wedge q) \rightarrow (Op \wedge Oq)$

(NEG-INFALLIBILITY) $O\neg Op \rightarrow \neg Op$

The knowledge operator satisfies NEG-INFALLIBILITY because knowledge is factive. But NEG-INFALLIBILITY also follows from the combination of:

(D) $Op \rightarrow \neg O\neg p$

(4) $Op \rightarrow OOp$

Proof: By 4, $Op \rightarrow OOp$. By D, $OOp \rightarrow \neg O\neg Op$. Putting these together, $Op \rightarrow \neg O\neg Op$. By contraposition, $O\neg Op \rightarrow \neg Op$.

Upshot: Any operator O that distributes over conjunction, if O is factive, or if O satisfies D and 4, the corresponding principles will be subject to the arguments here. The analogue of the knowledge principles for belief, justification, rationality, and evidence, insofar as these operators obey DISTRIBUTION and NEG-INFALLIBILITY, are also within the scope of this Fitch-style argument.

5. A Puzzle?

How did we get here? In the verificationism case, we started out with SVER which is obviously unpalatable, then moved to WVER. But that entails SVER, so this move is unavailable. In the closure case, we started out with SCLO which also (at least) appears unpalatable, then moved to WCLO. But that entails that it's impossible to know $p \wedge \neg Kq$ whenever p entails q, so this move also appears unavailable. Nor does the retreat to GLCO fare better.

Option 1: Reject one of the premises for Fitch's Paradox. For obvious reasons this blocks all such arguments. But what should we give up? Factivity? Distribution? Classical Logic?

There is a large literature on this. See e.g. the collection of essays in Salerno 2009. For a defense, see in particular Williamson 2000.

Option 2: Accept SCLO. In the verificationism case, SVER is very, very bad. How bad is SCLO? Maybe not as bad. In fact, if propositions are coarse-grained sets of possible worlds, DISTRIBUTION entails SCLO. So maybe let's just go with SCLO (and KK, EQUIV...) and explain away the putative counterexamples.

As Soames 1987 shows, on certain views according to which propositions coarse-grained sets of possible worlds, assuming that entailment is captured by the subset relation, if knowledge distributes over conjunction, then SCLO holds. *Proof.* Notice that $p \subseteq q$ iff $p \wedge q = p$. For suppose $p \subseteq q$. Then $p \wedge q = p \cap q = p$. Now suppose $p \wedge q = p$. Then $p \cap q = p$, so $p \subseteq q$. But let $p \subseteq q$ and consider Kp. Since $p \subseteq q$ iff $p \wedge q = p$, it follows that $K(p \wedge q)$. Distributing, Kp and Kq, so in particular, Kq. See also especially Stalnaker 1984.

Option 3: Accept that no version of closure is true. This appears unattractive:

> ... the idea that no version of [the closure principle] is true strikes me, and many other philosophers, as one of the least plausible ideas to come down the philosophical pike in recent years... (Feldman 1995, p.478)

Of course, even if no version of closure is true, it's perfectly consistent to think that some version of it *typically* holds.

Notable closure-deniers include Dretske 1970 and Nozick 1981.

Lesson: Fitch's Paradox isn't just a puzzle for verificationism – the issues associated with it are of much more general importance.

Appendix

We will work with a propositional language extended with two unary modal operators, K and $\Box$. We abbreviate n iterations of K as K^n and $\neg\Box\neg$ as $\Diamond$. A modal logic is a set of sentences which contains all tautologies of propositional logic and is closed under modus ponens and uniform substitution. We introduce the following axiom schemas and rule of inference:

For background, see for instance Chellas 1980.

(NEGATIVE-INFALLIBILITY$_K$) $K\neg Kp \rightarrow \neg Kp$

(DISTRIBUTION$_K$) $K(p \wedge q) \rightarrow (Kp \wedge Kq)$

(K$_\Box$) $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$

(4$_\Box$) $\Box p \rightarrow \Box\Box p$

(RN$_\Box$) From p, infer $\Box p$.

We will let L contain NEGATIVE-INFALLIBILITY$_K$, DISTRIBUTION$_K$, and K$_\Box$ and be closed under RN$_\Box$. In the appendix we prove some relevant facts about L that are referenced in the main text.

Fact 1. If L contains $p \to \Diamond Kp$, then L contains $p \to Kp$.

Proof. Suppose L contains $p \to \Diamond Kp$. By uniform substitution, $(p \wedge \neg Kp) \to \Diamond K(p \wedge \neg Kp)$. Assume, for contradiction, $K(p \wedge \neg Kp)$. Then by DISTRIBUTION$_K$, $Kp \wedge K\neg Kp$, from which $Kp \wedge \neg Kp$ follows by NEGATIVE-INFALLIBILITY$_K$. But $\neg(Kp \wedge \neg Kp)$, so $\neg K(p \wedge \neg Kp)$. By RN$_\Box$, $\neg\Diamond K(p \wedge \neg Kp)$. Therefore $\neg(p \wedge \neg Kp)$, which is equivalent to $p \to Kp$.

Fact 2. If L contains $p \to \Diamond Kp$, then L contains $p \to K^n p$.

Proof. We proceed by induction. For the base case, let $n = 1$. Then L contains $p \to Kp$ by Fact 1. For the inductive step, suppose L contains $p \to K^i p$. This is equivalent to $\neg(p \wedge \neg K^i p)$. By uniform substitution, $\neg(Kp \wedge \neg K^{i+1} p)$. By NEGATIVE-INFALLIBILITY$_K$, $\neg(Kp \wedge K\neg K^{i+1} p)$ and by DISTRIBUTION$_K$, $\neg K(p \wedge \neg K^{i+1} p)$. By RN$_\Box$, $\neg\Diamond K(p \wedge \neg K^{i+1} p)$ so $\neg(p \wedge \neg K^{i+1} p)$, which is equivalent to $p \to K^{i+1} p$.

SCLO and WCLO can be captured by the rules of inference:

(RM$_K$) From $p \to q$, infer $Kp \to Kq$.

(RM$_K^\Diamond$) From $p \to q$, infer $Kp \to \Diamond Kq$.

Hintikka (1970, p.142) writes that RM$_K$ is "apparently difficult to avoid in one's epistemic logic" while Rosenkranz (2021, p.58), writes that RM$_K$ "should be taken to fail" because of its strength.

Fact 3. If L is closed under RM$_K^\Diamond$ but not RM$_K$, then L contains $\neg\Diamond K(p \wedge \neg Kq)$.

Proof. Consider a pair of propositions p and q such that L contains $p \to q$ but not $Kp \to Kq$. Then L contains $(p \wedge \neg Kq) \to (q \wedge \neg Kq)$. By RM$_K^\Diamond$, $K(p \wedge \neg Kq) \to \Diamond K(q \wedge \neg Kq)$. But $\neg\Diamond K(q \wedge \neg Kq)$. Therefore L contains $\neg K(p \wedge \neg Kq)$, so by RN$_\Box$, L contains $\neg\Diamond K(p \wedge \neg Kq)$.

GCLO can be captured by he following rule of inference, where X is some set of conditions:

(G-RM$_K^\Diamond$) From $p \to q$, infer $(Kp \wedge X) \to \Diamond Kq$.

Fact 4. If L does not contain $\neg\Diamond(Kp \wedge X)$, then L is closed under G-RM$_K^\Diamond$ iff L is closed under RM$_K^\Diamond$.

Proof. The right-to-left direction is trivial. We prove the left-to-right direction. Let L contain $p \to q$, and suppose L is closed under G-RM$_K^\Diamond$ but not RM$_K^\Diamond$. Then Kp but $\neg\Diamond Kq$ for some p and q. But by G-RM$_K^\Diamond$, L contains $(Kp \wedge X) \to \Diamond Kq$, which is equivalent to $\neg\Diamond Kq \to \neg(Kp \wedge X)$ so by RN$_\Box$ and K$_\Box$, L contains $\Box\neg\Diamond Kq \to \Box\neg(Kp \wedge X)$. Since $\neg\Diamond Kq$ is equivalent to $\Box\neg Kq$, by 4$_\Box$, $\Box\Box\neg Kq$ and therefore $\Box\neg\Diamond Kq$. Therefore $\Box\neg(Kp \wedge X)$, which is equivalent to $\neg\Diamond(Kp \wedge X)$.

References

Chellas, B. 1980. *Modal Logic*. Cambridge: Cambridge University Press.

Das, N. and B. Salow. 2018. Transparency and the KK Principle. *Noûs*, 52, pp. 3-23.

David, M. and T. Warfield. 2008. In *Epistemology: New Essays*, ed. Q. Smith. Oxford: Oxford University Press.

Douven, I. 2005. A Principled Solution to Fitch's Paradox. *Erkenntnis*, 62, pp. 47-69.

Dretske, F. 1970. Epistemic Operators. *Journal of Philosophy*, 67, pp. 1007-1023.

Dretske, F. 2005. The Case Against Closure. In Matthias Steup and Ernest Sosa (eds.), *Contemporary Debates in Epistemology*, pp. 13-25. Blackwell.

Dummett, M. 2001. Victor's Error. *Analysis*, 61, pp. 1-2.

Feldman, R. 1995. In Defence of Closure. *The Philosophical Quarterly*, 45, pp. 487-494.

Fitch, F. 1963. A Logical Analysis of Some Value Concepts. *The Journal of Symbolic Logic*, 28, pp. 135-142.

Fraser, R. and J. Hawthorne. 2015. Cretan Deductions. *Philosophical Perspectives*, 29, pp. 163-178.

Greco, D. 2014. Could KK Be Ok? *The Journal of Philosophy*, 111, pp. 169-197.

Hart, W. D. 1979. The Epistemology of Abstract Objects: Access and Inference. *Proceedings of the Aristotelian Society*, 53, pp. 153-165.

Hawthorne, J. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.

Hintikka, J. 1970. 'Knowing that One Knows' Reviewed. *Synthese*, 21, pp. 141-162.

Immerman, D. 2020. Williamson, Closure, and KK *Synthese*, 197, pp. 3349-3373.

Liu, S. 2020. (Un)knowability and Knowledge Iteration. *Analysis*, 80, pp. 474-486.

McHugh, C. 2010. Self-knowledge and the KK Principle. *Synthese*, 173, pp. 231–257.

Nozick, R. 1981. *Philosophical Explanations*. Cambridge: Harvard University Press.

Salerno, J., ed. 2009. *New Essays on the Knowability Paradox*. Oxford: Oxford University Press.

San, W. K. 2020. Fitch's Paradox and Level-Bridging Principles. *The Journal of Philosophy*, 117, pp. 5-29.

Salow, B. manuscript. Fitch, KK, and Closure – a Reply to Liu.

Schaffer, J. 2007. Closure, Contrast, and Answer. *Philosophical Studies*, 133, pp. 233-255.

Soames, S. 1987. Direct Reference, Propositional Attitudes, and Semantic Content. *Philosophical Topics*, 15, pp. 47-87.

Stalnaker, R. 1984. *Inquiry*. Cambridge: Cambridge University Press.

Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.