

Vagueness and Introspection

Denis Bonnay & Paul Égré

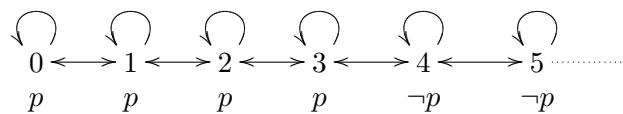
(Extended abstract)

Intransitivity and introspection. One central and debated aspect of the notion of inexact knowledge concerns the non-transitivity of the relation of indiscriminability and how it should be represented. On the epistemic account of vagueness put forward by Williamson, the intransitivity of the relation of indiscriminability is presented as the main source for vagueness ([5]: 237). In [4] and in the appendix to [5], Williamson formulates a *fixed margin for error* semantics for propositional modal logic in which the relation of epistemic uncertainty, based on a metric between worlds, is thus reflexive and symmetric, but non-transitive and non-euclidian. An important consequence of the semantics is that it invalidates the principles of positive introspection (if I know p , then I know that I know p) as well as negative introspection (if I don't know p , then I know that I don't know p).

In [1], we argued against Williamson that models of inexact knowledge that preserve the introspection principles can sometimes be desirable, and we presented a non-standard epistemic semantics for the notion of inexact knowledge, in which non-transitive and non-euclidian Kripke models can nevertheless validate positive as well as negative introspection. In [2], Halpern also argued against Williamson that an adequate model of vague knowledge need not invalidate the introspection principles, but following a different route. Instead of taking intransitivity as a primitive, and proving that the introspection principles can be preserved for a logic with one epistemic operator, as we did in [1], Halpern proposes a bimodal account of inexact knowledge that preserves the introspection principles, and he shows that there is a way to derive intransitivity. For Halpern, the intransitivity of vague knowledge is more characteristic of our reports on what we perceive than about our actual perception.

Despite these differences, one can establish a precise correspondence between Halpern's semantics and the semantics presented in [1]. The object of this paper is to spell out the details of this correspondence, and thus to compare two strategies in order to keep together introspection and non-transitivity. Like Halpern, but contra Williamson, we think it does make sense to preserve the introspection principles within a logic of inexact knowledge; unlike Halpern, but in agreement with Williamson, we are ready to see non-transitivity as a property of perceptual knowledge proper.

Centered Semantics. Consider a discrete series of pens linearly ordered by size, such that all and only pens that are less or equal than 4 inches fit in a certain box. A subject sees the pens and the box at a certain distance and is asked which pens will fit in the box. Let us suppose that the subject cannot perceptually discriminate between pairwise adjacent pens, namely between pens whose size differs by less than 1 cm. The scenario may be represented by means of the following Kripke structure, in which p represents the objective property of fitting in the box, with worlds indexed by sizes.



With respect to that model, $2 \models \Box p$, and $3 \models \neg\Box p \wedge \neg\Box\neg p$: thus, the subject knows that an object of size 2 will fit in the box, and does not know whether an object of size 3 fits in the box. Crucially, $2 \models \neg\Box\Box p$, that it is the subject doesn't know that he knows that the objects fits in the box, since 4 is a $\neg p$ -world accessible in two steps from 2. For Williamson, this result is welcome, since iterations of knowledge operators are seen as a “process of gradual erosion” in the case of vague knowledge. However, one may argue that, looking at a pen of size 2, my knowing that I know that it will fit in the box supervenes only on my knowing whether it fits in the box, and not on epistemic alternatives that are too far. In [1], we formulated an alternative semantics (CS, for centered semantics), in which the epistemic alternatives relevant for iterated modalities remain the worlds accessible in one transition from the world of evaluation. In other words, every fact concerning the knowledge of the agent should be decided solely on the basis of worlds that are not distinguishable from that world, without having to move further along the accessibility relation. Given a model $M = \langle W, R, V \rangle$, we first define the notion of satisfaction for couples of worlds, and extract the definition of satisfaction for single worlds:

Definition 1. *CS-satisfaction for couples of worlds:*

- (i) $M, (w, w') \models_{CS} p$ iff $w' \in V(p)$.
- (ii) $M, (w, w') \models_{CS} \neg\phi$ iff $M, (w, w') \not\models_{CS} \phi$.
- (iii) $M, (w, w') \models_{CS} (\phi \wedge \psi)$ iff $M, (w, w') \models_{CS} \phi$ and $M, (w, w') \models_{CS} \psi$.
- (iv) $M, (w, w') \models_{CS} \Box\phi$ iff for all w'' such that wRw'' , $M, (w, w'') \models_{CS} \phi$.

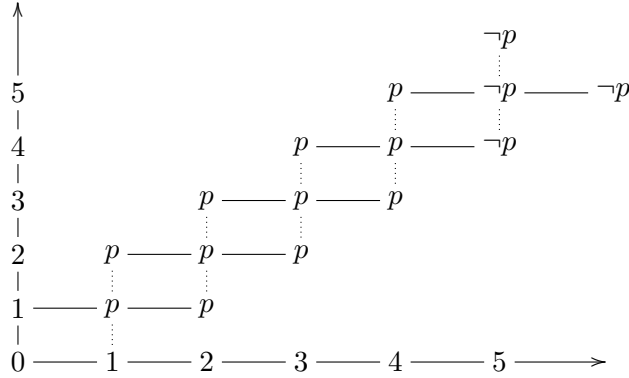
Definition 2. $M, w \models_{CS} \phi$ iff $M, (w, w) \models_{CS} \phi$

Clause (iv) of the definition accounts for the “centered” feature of the semantics, for it entails that for every w and w' : $M, (w, w') \models_{CS} \Box\phi$ iff $M, (w, w) \models_{CS} \Box\phi$ iff $M, w \models_{CS} \Box\phi$. This ensures that instead of looking at worlds that are two steps away to check whether $\Box\Box\phi$ is satisfied, one backtracks to the actual world to see whether $\Box\phi$ already holds there. In the previous model, it can be checked that $2 \models_{CS} \Box p$, and likewise $3 \models_{CS} \neg\Box p \wedge \neg\Box\neg p$. However, $2 \models_{CS} \Box\Box p$ and $3 \models_{CS} \Box\neg\Box p$. In [1], we proved that the normal logic K45 is indeed sound and complete with respect to CS. Likewise, one can formulate a centered version of Williamson’s fixed-margin semantics, CMS, for which S5 is sound and complete.

When \Box is interpreted as “it is clear that”, it may be objected that CS makes room only for first-order vagueness, and not for higher-order vagueness, since “it is not clear that p ” entails “it is clear that it is not clear that p ”. To this, however, two replies can be made: firstly, when \Box is read as “I know that”, as we assumed, $\Box\neg\Box p$ should rather mean that I am aware of my uncertainty at the moment it first arises: by analogy to situations of “forced march”, in which I am forced to answer by “yes” or “no”, this means that I am aware of making a “jump” in my judgements when the jump occurs (see e.g. [3]); moreover, we could imagine that the fluctuation characteristic of higher-order vagueness is caused by a fluctuation of amplitude in the margin of error over time, or by other factors. Secondly, we show in [1] that CS is a particular case of a family of parameterized semantics $TS(n)$ (for “token semantics”), for which the trivialization of the iterations need not occur at the first level, but at any arbitrary level n of iterated modalities, depending on the number n of tokens available (the intuition is that moving along the accessibility relation has a cognitive cost, which is mirrored by the fact that a token has to spent for each move in a model, and the number of tokens available at the beginning is finite). Thus, $TS(n)$ and standard Kripke semantics coincide for formulas with less than n embedded modalities.

Halpern’s semantics. Halpern takes a different approach to this problem, since his logic allows to get distinct syntactic representations of the operators “I know that” and “it is clear that”. His logic

(in the one-agent case) has two primitive operators, namely R and D , where $R\phi$ means that the agent “reports ϕ ”, and $D\phi$ means that “according to the agent, ϕ is definitely the case”. A model, relative to this language, is a structure $\langle W, P, \sim_s, \sim_o, V \rangle$, where $W \subseteq S \times O$, where S is intended to denote a set of subjective states and O a set of objective states. The relations \sim_o and \sim_s both are equivalence relations over W , and V is a valuation over W . P , finally, is a subset of W , intended to denote the states that the agent considers plausible. For simplicity, we shall assume that $P = W$ here, and therefore we shall omit reference to P in the definition of satisfaction. Modulo that simplification, the satisfaction clauses for the modal operators are the expected ones, namely $M, w \models R\phi$ iff for every w' such that $w \sim_s w'$, $M, w' \models \phi$, and similarly for $D\phi$ with respect to \sim_o . As a consequence, each operator is axiomatized by the logic S5.¹ The point of Halpern’s approach, however, is that although each operator separately obeys transitivity (and euclidianity), their combination DR need not (if two binary relations A and B are equivalence relations, it does not follow that their composition $A \circ B$ is transitive or euclidian). Intuitively, an agent definitely reports that ϕ when his estimation is sufficiently reliable, just as in Williamson’s approach. In this way, the complex operator (DR) plays exactly the role of Williamson’s “clearly” operator in margin for error semantics.



To make the link concrete, let us consider a model in which W is the subset of $\mathbb{N} \times \mathbb{N}$ consisting of couples (n, m) such that $|n - m| \leq 1$. Let us suppose that n is the objective size of some object, or the objective value of some parameter, and m its subjective estimate. Let us suppose moreover that $(n, m) \sim_o (n', m)$ iff $n = n'$, and likewise $(n, m) \sim_s (n', m')$ iff $m = m'$. It is easy to check that both relations are equivalence relations over W . In the above figure, each cell of the partition determined by \sim_o corresponds to the points connected by a vertical dotted line, and each cell of the partition determined by \sim_s corresponds to the points connected by a horizontal straight line. Let us suppose moreover that whether a point w is a member of $V(p)$ depends only on the objective part of w . For instance suppose that $(n, m) \in V(p)$ iff $n < 5$ (as in our previous example, p may stand for “fitting in the box”). It is easily checked that $(2, 3) \models DRp$, but $(2, 3) \not\models DRDRp$. Thus, if the size of the object is 2 and the measurement of the agent is 3, with a threshold for $\neg p$ that is between 4 and 5, then the agent definitely reports that p , but will not iterate this judgement. By contrast, R is an S5 modality, satisfying negative and positive introspection at any point in the model.

In the previous model, Dp is equivalent to p . Thus, if we consider only the relation of *subjective* equivalence for R , a model like the model of Figure 2 may be called a *layered margin model*, since each horizontal equivalence class (namely the classes for \sim_s) contains the possible objective values

¹In Halpern’s full version of the semantics for the multi-agent case, each modality is actually a KD45 operator, and the D_i satisfy a weakened version of axiom **T**.

that are compatible with the agent's subjective parameter, and the horizontal projection of these classes onto the x -axis of the model would yield a linear structure of inexact knowledge of exactly the kind with which we started. This notion of layering can be made precise. Thus, given a Kripke model $M = \langle W, R, V \rangle$, let us call $L(M) = \langle W', R', V' \rangle$ a *layering* of M if it satisfies: $W' = \{(w, w') \in W \times W; w' R w \vee w' = w\}$; $(w, w') R'(u, u')$ iff $w' = u'$ and $w' R u$; and finally, $(w, w') \in V'(p)$ iff $w \in V(p)$. It can be checked that R' in $L(M)$ is necessarily transitive and euclidian. It is easy to establish that, relative to the basic modal language:

Proposition 1. $M, w \models_{\text{CS}} \phi \iff L(M), (w, w) \models \phi$

by proving that for all (w', w) in $L(M)$:

Lemma 1. $M, (w, w') \models_{\text{CS}} \phi \iff L(M), (w', w) \models \phi$.

Proof. By induction on ϕ .

□

Comparisons. Proposition 1 applies also to Williamsons's fixed margin models $M = \langle W, d, \alpha, V \rangle$ (see [4]) in which two worlds w, w' are accessible iff $d(w, w') \leq \alpha$, where d is a metric over W . With respect to margin models, Proposition 1 shows that Halpern's operator R therefore plays exactly the role of the knowledge operator \Box in the framework of centered semantics. On the one hand, these results makes clear that centered semantics really is a standard two-dimensional semantics in disguise. On the other hand, the operation of layering shows how it is possible to recover transitivity (and euclidianity) from a non-transitive (and non-euclidian) relation. In Halpern's approach, the intransitivity characteristic of qualitative comparison is simulated by means of two operators. Unlike Halpern, we don't think that any epistemic uncertainty relation should necessarily be transitive. Fundamentally, however, the two approaches follow essentially the same inspiration, by making introspective knowledge depend only on the states that are subjectively accessible to the agent and by avoiding spurious dependencies to objective alternatives that are too far apart. Finally, the idea of layered models constructed out of non-transitive and non-euclidean models can be generalized to Token Semantics: n -layered models can be defined, on which the standard semantics satisfies the same formulas as does Token Semantics with n tokens on the ground model.

References

- [1] D. Bonnay & P. Égré (2006), A non-standard Semantics for Inexact Knowledge with Introspection, forthcoming in S. Artemov & R. Parikh (eds), *Proceedings of the ESSLLI 2006 Workshop on Rationality and Knowledge*, Malaga.
- [2] J. Halpern (2004), Intransitivity and Vagueness, *Proceedings of the Ninth International Conference on Principles of Knowledge Representation and Reasoning (KR 2004)*, pp. 121-129.
- [3] S. Shapiro (2003), Vagueness and Conversation, in J.C. Beall (ed.), *Liars and Heaps*, pp. 39-72, Oxford UP.
- [4] Williamson T. (1992), Inexact Knowledge, *Mind*, 101, pp. 217-42.
- [5] Williamson T. (1994), *Vagueness*, Routledge.