THROUGH THE LOGICIAN'S GLASS

HYKEL HOSNI

Abstract. This tutorial assumes no previous knowledge of logic. It only assumes that readers have an interest in some key questions that arise in the social sciences. Logicians who have never been exposed to such questions will hopefully appreciate how far beyond the usual domains of application logic can stretch. The tutorial's main purpose is therefore to encourage cross-fertilisation by highlighting how logic provides a flexible analytical tool to the social sciences and how the social sciences abound with pressing problems of great potential interest to logicians.

Contents

La logica è come un palo, utile in quanto può impedire alla pianta del pensiero di crescere storta. Ma, come un palo non è una pianta né il possibile surrogato di una pianta, così la logica non è il pensiero né una specie di surrogato del pensiero. (Bruno de Finetti)

Date: July 8, 2013.

Tutorial for *Games and Decisions*, Centro di Ricerca Matematica "E. De Giorgi", Pisa 8-10 July 2013. http://crm.sns.it/event/275/ .

1. Three questions on rationality

According to a standard view, decision theory divides into three main branches, namely *individual* decision theory (sometimes called decision theory *tout court*) *interactive* decision theory (aka game theory) and *social choice* theory. Though the overlaps are significant, the tripartition is rooted in three distinct questions, respectively:

(1) what are the "rationality" constraints to be imposed on an individual who must select one from a set of alternative courses of action whose consequences are uncertain?
(2) what is a "rational" solution to an interaction among agents who are rational in the sense spelled out in (1)?
(3) how are we to aggregate individually "rational" preferences into a collective profile which expresses the preference of society as a whole?

This tutorial aims at giving readers with no previous knowledge of the subject a glimpse into how logic can provide an interesting perspective on the above questions. The tutorial is organised as follows. Section 2 introduces the key concept of *compositionality* which will lead us to put a central problem of *aggregation* in a very simple logical form. We then move on to what is arguably the single most important notion in the vast logical landscape, the formal notion of *consequence*. In Section 3 I will illustrate how logical consequence plays a fundamental role in constraining probability functions, thereby unravelling an interesting perspective on how logic relates to decision theory under uncertainty. Section 4 focusses on a methodological feature of logic, which I will call Axioms-As-Properties, or AxAP, for short. As an illustration of its applicability I will introduce some basic notions from *epistemic logics*, i.e. those logics which aim at providing a rigorous definition of "knowledge" which lies at the very heart of solution concepts in game theory.

The main goal of this note is therefore to suggest that logic can provide exciting perspectives on central problems in each of the three main branches of decision theory. In the limited space of this note I can only hope to say enough to get readers interested in finding out more. Each Section ends with a selection of references which will hopefully give interested readers enough guidance to pursue their interests.

Before getting started, let me say something about logic, in general.

1.1. **Logic.** Logicians with a taste for applying their work to analysing rationality[1] often assume that drawing consequences from available premisses is all there is to rationality. Of course they do not mean this in a literal way. It is rather a commonly accepted *abstraction* aimed at stripping away the inessential aspects which accompany the manifestations of rationality from which logicians draw their inspiration.

A paramount example of this is provided by Alan M Turing's (1912 - 1954) analysis of computation. Standing on the shoulders of Leibniz and all his less well-known forerunners, Turing isolated the essential aspects of the human activity of *computing*. This led him to a mathematical definition of computation which dispensed from the requirement to the effect that the computor be a human being. The result is what we now call a Turing Machine, which lies at the very root of the digital information age.

There is at present no reason to believe that modelling what we refer to as "intelligence" should require us to go beyond handling strings of binary digits. However, logicians find it useful to tackle the problem of rational reasoning by looking at which properties are satisfied by certain *consequence relations*. By doing this logicians make (albeit often only implicitly) the sort of abstraction pioneered by Turing, and identify the "rationality" of an agent with a set of desirable properties of a consequence relation. This allows logicians to set otherwise hard-to-pose questions on the nature of rationality in a terse mathematical language.

What follows is a short guided tour of how this happens and of its relevance to decision theory broadly conceived.

2. Compositionality

It is a good approximation to say that *sentences* are characterised by the fact that it makes sense to ask whether they are true or false. Consider the following

(1) The Scuola Normale Superiore is in Pisa
(2) The Scuola Normale Superiore is in Milan
(3) Where is the Scuola Normale Superiore?
(4) The President of the United Kindgom is a former leader of the Labour Party

Whilst (1) and (2) are easily settled (3) and (4) do not qualify as sentences for our purposes, though they make sense, to some degree.

[1]To whom I will refer simply as *logicians* in the remainder of this note.

We start with a language $\mathcal{L}$, which for our purposes is going to be a finite set of propositional variables $\mathcal{L} = \{p_1, p_2, \ldots, p_n\}$. Intuitively, elements of $\mathcal{L}$ are thought of as the smallest units for which it makes sense to ask whether they are true or false. Those building blocks can then be used to form increasingly more complex sentences by suitably applying four *propositional connectives*, namely $\neg, \wedge, \vee, \rightarrow$ which read as "negation", "conjunction", "disjunction" and "implication", respectively. Connectives are used to combine the (finite) set of propositional variables in $\mathcal{L}$ to give rise to the *infinite* set of *sentences of* $\mathcal{L}$, denoted by $\mathcal{SL}$, as follows. We start by saying that all propositional variables are sentences, i.e.

$$\mathcal{SL}_0 = \mathcal{L}.$$

Suppose now that we have combined several simpler sentences to give rise to more complicated ones, and have done so n times. The following condition says how we may be able to get from n to $n+1$:

$$\mathcal{SL}_{n+1} = \mathcal{SL}_n \cup \{\neg\theta, (\theta * \phi) \mid \theta, \phi \in \mathcal{SL}_n, * \in \{\wedge, \vee, \rightarrow\}\}$$

Note that n is arbitrary here. Finally we add a condition that says that we can iterate this (recursive) construction for as long as we please, and by doing so we get an infinite set out of the combination of a finite number of building blocks ($\mathcal{SL}_0$) and four connectives:

$$\mathcal{SL} = \bigcup_{n \in \mathbb{N}} \mathcal{SL}_n.$$

[*Aside.* A kind of question which typically catches the logician's fancy is what is the smallest set of propositional connectives which is sufficient to give rise to the whole of $\mathcal{SL}$.]

It fits our purposes to say that "θ holds" means that the agent whose reasoning we are modelling takes θ to be true. We make this precise by defining propositional *valuations* as functions

$$v : \mathcal{L} \rightarrow \{0, 1\}.$$

Logicians read $v(\theta) = 1$ as "θ is true", and refer to "1" as θ's truth-value.

Valuations allow us to provide meaning to our propositional connectives by means of the following *truth table* pictured in Table 1

It is an easy but instructive exercise to show that valuations extend uniquely to $\mathcal{SL}$. So truth-tables analogous to Table 1 can be written for arbitrarily complex (finite) sentences. This suggests that the truth-value of, say a conjunction is a fixed function of the truth-value of the conjuncts. More precisely, let $\theta, \phi \in \mathcal{SL}$, then there exists a function

$$f_\wedge : \{0, 1\} \times \{0, 1\} \rightarrow \{0, 1\}$$

p	q	$\neg p$	$p \wedge q$	$p \vee q$	$p \to q$
1	1	0	1	1	1
1	0	0	0	1	0
0	1	1	0	1	1
0	0	1	0	0	1

TABLE 1. The meaning of propositional connectives

such that

$$v(\theta \wedge \phi) = f_{\wedge}(v(\theta), v(\phi)).$$

The reader is encouraged to find a function that satisfies the above, and to do the same for each connective. The (central!) idea that the truth-value of (arbitrarily complex) sentences is a fixed-function of the truth-value of their components is known as the *principle of compositionality*.

2.1. **Coherence, consistency and satisfiability.** Let $\mathcal{L} = \{p, q, r\}$. Suppose three jurors must decide whether a defendant is liable (r). According to contract law, a defendant is liable if and only if they are under the obligations of a valid contract (p) *and* they are in breach of it (q).

The so-called *doctrinal paradox*, arises when Jurors submit the following judgments:

	p	q	r
Juror A	1	1	1
Juror B	1	0	0
Juror C	0	1	0
Majority	**1**	**1**	**0**

Contract law says that $(r \to (p \wedge q) \wedge (p \wedge q) \to r)$, and it is implicitly assumed that all Jurors accept that. Under this assumption, the of the above becomes apparent. Whilst individually each Juror judges in accord with the constrains imposed by the definition of propositional connectives as presented in Table 1, the simple-majority aggregation of their judgments by simple majority gives does violate such constraints. Indeed the Majority's judgment is *inconsistent*.

Inconsistency is a central notion in logic. To some extent logic claims its normative role by freeing reasoners who conform to it from inconsistency. Hence, many would regard *consistency* (or coherence) to be the single most important contribution made by logic to the normative analysis of rationality. We approach the formal definition of (in)consistency via the notion of *satisfiability*.

Let v be a propositional valuation on $\mathcal{L}$. We say that v *is a model of* $\Gamma \subseteq \mathcal{SL}$ if $v(\gamma) = 1$ for all $\gamma \in \Gamma$ (which we abbreviate by writing $v(\Gamma) = 1$). We say that Γ is *satisfiable* if it has a model, and *unsatisfiable* otherwise.

This captures a central feature of coherence. For if a set is unsatisfiable, there is no way in which its elements can be all true, that is to say they do not *cohere*.

It is a simple exercise to show that taken individually all jurors submit satisfiable judgments, but that the simple-majority aggregation of their individual judgments is unsatisfiable. Hence incoherent. If coherence is identified with rationality (and at this level of abstraction, this seems entirely plausible), the discursive dilemma is a situation in which *individual rationality* leads, via simple majority aggregation, to *collective irrationality*.

There is more than analogy between this and the celebrated Arrow's impossibility theorem. Building on an elementary logical setting of the sort just outlined Dietrich and List (2007) prove an impossibility result for judgment aggregation from which Arrows' impossibility theorem is derived as a corollary. Hence aggregating preferences is less fundamental than aggregating judgments. This is just one example of how the logical analysis of aggregation may shed interesting light on this central problem.

2.2. **Further reading.** The idea that logic is to do with the formalisation of coherence is articulated very accessibly in Hodges (1985). I wholeheartedly recommend this as an entry point to the subject.

The area of Judgment aggregation originates with a problem known in legal theory as the *doctrinal paradox* (Kornhauser and Sager, 1986), but entered the social choice literature through (Pettit, 2001), who termed it *discursive dilemma*. Despite its very recent history, there is a very substantial body of work on judgment aggregation, which is very well documented by List (2011). (Williamson, 2010, Chapter 8) puts forward a model of judgment aggregation based on merging *evidence* rather than judgments directly and discusses an application of this to (medical) decision-making.

3. Consequence

If we identify rationality with coherence, the mathematical definition of satisfiability can be used to define what it means to reason *irrationally*, namely to move from premisses which are satisfied to conclusions which are not. Since not all sets are satisfiable, satisfiability can be seen as an asset which we certainly don't want rational reasoning to waste. Hence the central feature of logical consequence is that it should *preserve satisfaction.*

Definition 3.1 (Tarski 1936)**.** θ is a logical consequence of Γ if and only if every model of Γ is a model of θ, which we write as follows

$$\Gamma \models \theta \Leftrightarrow M_\Gamma \subseteq M_\theta.$$

Example 3.1. Let $\Gamma = \{p, p \to q\}$ and let $\theta = q$. Is it the case that $\Gamma \models \theta$? Compositionality allows us to answer in a way which logicians call *effective.* This amounts to say that we possess a procedure to answer definitely (either Yes or No) which is guaranteed to end after a finite number of steps. One such method is provided by truth-tables as follows:

p	q	$p \to q$	q
1	**1**	**1**	**1**
1	0	0	0
0	**1**	**1**	**1**
0	0	1	0

Each row corresponds one of the 4 possible valuations on $\mathcal{L}$. By the Definition 3.1 we need to check whether there exists a model of Γ which *does not satisfy* θ. A quick inspection of the table shows that no such model exists, hence it is indeed the case that $\Gamma \models \theta$.

We say that a sentence θ is a *tautology*, and write $\models \theta$, if and only if every valuation satisfies θ (alternatively: θ is true under all the $2^{\mathcal{L}}$ valuations on $\mathcal{L}$, alternatively: if it is mapped to 1 in each row of a suitable truth-table). A tautology is therefore true "in virtue of its logical structure", and therefore we don't need to mention its models. This explains the notation. A *contradiction* is defined similarly, i.e. as a sentence which has no models. So contradictions are unsatisfiable.

One cannot overemphasise the *formal* nature of consequence. Whether a sentence θ is a logical consequence of some premisses Γ must not depend on the intuitive meaning which we associate to those sentences, but only on their logical form. The reader may get a glimpse of this by cooking up a valid relation $\Gamma \models \theta$ which nonetheless can be given an implausible rendering (by suitably interpreting the sentences). In addition, it is an instructive exercise to show that *any* sentence is a logical consequence of unsatisfiable premisses. Whether this property of consequence (often referred to as *ex falso quodlibet (sequitur)*) is acceptable or not for the normative science of reasoning, has bothered logicians for centuries.

3.1. **Consequence and probability.** Definition 3.1 supports the common view that logic is to do with *certainty*. And that's in fact how it started. Yet, since the mid-seventeenth century, the idea idea that logic had something important to say about the *ars inveniendi* became increasingly popular. In 1704 Leibniz put it as follows

> I maintain that the study of the degrees of probability would be very valuable and is still lacking, and that is a serious shortcoming in our treatises on logic. For when one cannot absolutely settle a question one could still establish the degree of likelihood on the evidence, and so one can judge rationally which side is the most plausible[2].

Probability logic, and the logical approach to probability, both address the *lacuna* lamented by Leibniz. The remainder of this section is devoted to illustrating how probability can be given a simple, yet extremely interesting, logical footing.

3.2. **A representation theorem for probability functions.** A probability function over the language $\mathcal{L}$ is a map $P : \mathcal{SL} \to [0,1]$ satisfying

(P1): If $\models \theta$ then $P(\theta) = 1$
(P2): If $\theta \models \neg\phi$ then $P(\theta \vee \phi) = P(\theta) + P(\phi)$.

It is interesting to note that a *finite* subset of $\mathcal{SL}$ carries all the relevant information about probability functions. To see this, we need a little notation. Let $p \in \mathcal{L}$ and let $p^1 = p$ and $p^0 = \neg p$. Then, for a valuation v and $\epsilon \in \{0,1\}$:

$$v(p^\epsilon) = 1 \Leftrightarrow \left\{ \begin{matrix} \epsilon = 1 \text{ and } v(p) = 1 \text{ or} \\ \epsilon = 0 \text{ and } v(\neg p) = 1 \end{matrix} \right\} \Leftrightarrow v(p) = \epsilon$$

Let $AT^{\mathcal{L}}$ be the set of atoms of $\mathcal{L}$, that is the set of sentences of the form

$$\alpha = p_1^{\epsilon_1} \wedge p_2^{\epsilon_2} \wedge \ldots \wedge p_n^{\epsilon_n},$$

where $\epsilon_i \in \{0,1\}, i = 1, \ldots, n$.

Notice that the set $AT^{\mathcal{L}}$ is in 1-1 correspondence with the valuations on $\mathcal{L}$. This implies that there is a unique valuation satisfying $v(\alpha) = 1$ namely $v_\alpha(p_i) = \epsilon_i$ for $1 \leq i \leq n$. Conversely, given a valuation $v \in \mathbb{V}$ there exists a unique atom $\alpha \in \mathcal{AT}^{\mathcal{L}}$ such that $v(\alpha) = 1$, namely that $\alpha = \bigwedge_{i=1}^n p_i^{\epsilon_i}$ for

[2] Leibniz *Nouveaux essais*, translated by P. Remnant e J. Bennet (1981) p. 372

which $\epsilon_i = v(p_i)$ for $1 \leq i \leq n$, i.e. the atom:

$$\alpha = \bigwedge_{i=1}^{n} p_i^{v(p_i)}.$$

Let

$$M_\theta = \{\alpha \in AT^{\mathcal{L}} \mid \alpha \models \theta\}.$$

Since there exists a unique valuation satisfying α, say v_α, by definition of $\models$ it must be the case that $v_\alpha(\theta) = 1$. Thus

$$M_\theta = \{\alpha \in AT^{\mathcal{L}} \mid v_\alpha(\theta) = 1\}.$$

Theorem 3.1 (Paris 1994)**.**

(1) *Let P be a probability function on $\mathcal{SL}$. Then the values of P are completely determined by the vales it takes on the atoms of $\mathcal{L}$, as fixed by the vector*

$$\langle P(\alpha_1), P(\alpha_2), \ldots, P(\alpha_J)\rangle \in \mathbb{D}^L = \{\vec{x} \in \mathbb{R}^J \mid \vec{x} \geq 0, \sum_{i=1}^{J} x_i = 1\}.$$

(2) *Conversely, fix $\vec{a} \in \mathbb{D}^L$ and let $P' : SL \to [0,1]$ be defined by*

$$P'(\theta) = \sum_{\alpha \in M_\theta} a_i.$$

Then P' is a probability function.

Theorem 3.1 brings clearly to the fore the heavy lifting done by logical consequence in constraining probability values. Suppose now that you felt dissatisfied with some properties of the relation $\models$, say the fact that

$$\models \theta \vee \neg\theta \tag{1}$$

also known as the "law of excluded middle". Well, if you do feel that, you are very likely to sympathise with an important research strand known as *intuitionistic logic*, which in turn constitutes the main foundational framework for the development of *constructive mathematics*. Suppose you reject, as intuitionistic logicians to, the validity (in general) of (1). Then you would very probably be dissatisfied with an immediate consequence of P1-P2 above, namely

$$P(\theta \vee \neg\theta) = 1 \tag{2}$$

aka *probabilistic excluded middle.* A common argument for rejecting the universal adequacy of (2) is that we might not know anything at all about θ, so why giving a statement about it *the highest* of degrees of belief? To some people this just doesn't seem rational – chances are they have strong constructive sympathies.

It might come as a surprise to non logicians that an beautiful and densely populated galaxy of non-classical logics exists, allowing many venues of non-classical probability to be explored. Indeed dissatisfaction with probability as a norm of rational belief under uncertainty may sometimes be better understood as dissatisfaction with $\models$ as the rational norm of reasoning under certainty.

Finally, the logical setting of theorem 3.1 offers an interesting take on the vexed question of the *state space* in decision theory. For in our setting controversies related to the determination of a suitable state-space can hardly arise. Once a language $\mathcal{L}$ has been fixed, there is a unique state space, namely the set of atoms $\mathcal{AT}^{\mathcal{L}}$ over which the unit mass is to be distributed. Pushing the problem of the state-space one level up to the choice of a logical language has interesting consequences. I'd like to mention briefly two. First, as noted in (Section 9.2 Williamson, 2010), language has an epistemic value, hence certain languages are naturally more suited to capture certain phenomena of interest than others. So, for example, Italian has generally been regarded as the operatic language *par excellence*, whereas German has been often recognised as a favourite language for writing mathematics. Second, it makes perfect sense to assume that one individual possesses a plurality of languages. Again, the choice of the most suitable language, and hence of state-space, may be constrained by context-dependent evidence. Put otherwise, the choice of a language appears to be significantly less arbitrary than the choice of a state-space.

3.3. **Further reading.** The notion of logical consequence has been around for as long as the idea of logic has, but it became mathematically precise only after the seminal work of Alfred Tarski (1901 - 1983). Definition 3.1 lies at the very heart of mathematical logic, for which many excellent introductions are available, including (Enderton, 2001; Boolos et al., 2002; Chiswell and Hodges, 2007). Readers with a good command of classical (propositional) logic, will find chapter 1 of Bochman (2001) a concise and elegant introduction to Tarskian consequence operators.

Probability logic is a rapidly growing field which is being fed by parallel and research strands. A very good introduction to the area is provided by (Haenni et al., 2011), which assumes very little. Mathematically inclined readers with a good command of classical first-order logic are suggested to engage with (Paris and Vencovska, 2013). This book extends to first-order logic the propositional setting of Paris (1994). Finally, (Jeffrey, 2004; Adams, 1996) constitute exciting entry points for philosophically minded readers.

The logical representation of probability function is made possible by the algebraic properties of propositional logic. Readers who wish to explore how propositional connectives act as operations on truth-values to give rise to boolean algebras are highly recommended to start with (Halmos and Givant, 1998).

The investigation on probability based on non-classical logics is in its infancy. One currently very active research strand started with (Paris, 2001) and is recapped in the introductory section of (Fedel et al., 2011). Although the paper doesn't address the question of adding a (probability) measure on the algebra, I'd like to mention Ciraulo et al. (2013) which explores constructive versions of boolean algebras.

4. Axioms As Properties (AxAP)

Whereas (individual) decision theory focusses on understanding uncertainty and ignorance, i.e. what agent's don't know, game theory has a distinctive interest on what agents *do know*, and what they known about who knows what. In this context knowledge is essential in reducing *strategic uncertainty.* The simplest example being the role of the assumption of *common knowledge of rationality* in the elimination of dominated strategies.

For a relatively long time, economic theorists relied on their intuition in relation to the notion of "knowledge", until Aumann (1976) put forward what have become known as *Aumann structures.* But do we really know what we are talking about when reasoning about knowledge? According to standard *epistemic logic* we should know – that's called *positive introspection* in the trade.

Interestingly enough the standard epistemic logic, known as S5, turns out to be isomorphic to Aumann's structure. Hence, by and large, epistemic logicians (philosophers, mathematicians and computer scientists) agree with economists on what should count as knowledge. This is all the more surprising given that game theorists and logicians ignored for a long time each other's work.

This tutorial ends by illustrating the remarkable flexibility licensed by the logical formalisation of knowledge.

4.1. **Aiming at a normatively adequate formalisation of knowledge.** Epistemic logic had an interesting feedback on the pre-formal conception of knowledge, or more precisely on what we should take as a normatively adequate notion of knowledge. This is one of two secrets for the success of epistemic logics in epistemology and artificial intelligence. The way epistemic logic could feed back on purely epistemological investigation on the nature of knowledge depends on a key feature of logical formalisation which I will refer to as Axioms As Properties (AxAP). The idea, quite simply, is that the notion of knowledge is too complex to be analysed directly, so it is broken down to properties which are considered to be *necessary* for any reasonable formalisation of knowledge to satisfy. As usual with mathematical axiomatisation, it is a lot easier to list necessary conditions and claim they're also sufficient when we don't seem to be able to list more. So, for instance, many epistemologists believe that it is necessary to any normatively adequate account of knowledge that we should not be in a position to know falsehoods, as captured by Axiom T below. The beauty of the AxAP approach is that you can judge your axioms independently of one another. Thus, a favourite logic for a "revisable" notion of knowledge, which many find more suited to scientific (empirical) knowledge, is obtained by replacing T with one which intuitively requires knowledge to be *consistent* rather than true.

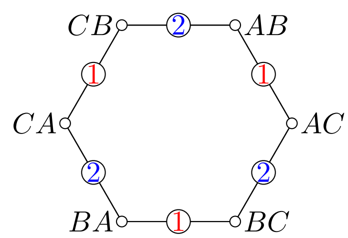
The second reason for the success of epistemic logic is again to do with correspondence, but this time of a *formal* kind. The idea is that each axiom singles out a class of *graphs* that serve as a semantics for the corresponding logic. To appreciate this we have to introduce some more logical notions and notation.

Let N be a set of agents, S be a set of states (often deceivingly called "possible worlds") and $\theta, \phi \in \mathcal{SL}$. The key intuition motivating *epistemic logics* is that agent i, whom we assume to be in epistemic state s, may *access* other epistemic states. This motivates the introduction of an *accessibility* relation $R_i \subseteq S^2$ (one for each agent $i \in N$). We say that an agent $i \in N$ in epistemic state $s \in S$ *knows* θ just if θ is true (as above) *in all epistemic states* which i can access from s. Distinct formalizations of "knowledge" (and "belief") arise by imposing distinct constraints on the accessibility relation R. In the most widely studied (multi-agent) epistemic logic, known as $S5$, R is an *equivalence* relation (reflexive, symmetric and transitive), thus making the resulting logical characterization of knowledge effectively equivalent to that provided by Aumann's structures. This is all rather abstract, so let us move on to an example which illustrates well the appropriateness of the equivalence relation.

Example 4.1 ((Fagin et al., 1996)). Consider two players $N = \{1, 2\}$ and a deck of three (distinct) cards labelled A, B, C. Each player is dealt a card, whilst the third card is left face down on the table. There are six possible configurations which constitute meaningful epistemic states

$$S = \{AB, AC, BC, BA, CA, CB\},$$

where s_1 reads as "player 1 is dealt card A, player 2 is dealt card B and card C is left on the table" etc. The idea here is that in state $s = (x, y)$ player 1 has access to (hence considers possible) states $s = (x, y)$ and $t = (x, z)$ whereas player 2 has access to (hence consider possible) states $s = (x, y)$ and $u = (z, y)$, where of course $x \neq y \neq z$. This gives rise to the following *graph representation* of the situation, where the edges between states are labelled with the players' names:



This can be generalised by introducing a little logical background and notation. First we need to extend the set $\mathcal{SL}$ by closing under the epistemic operator $K(\cdot)$ as follows:

$$\begin{aligned}
\mathcal{ESL}_0 &= \mathcal{L} \\
\mathcal{ESL}_{n+1} &= \mathcal{ESL}_n \cup \{\neg\theta, K_i\theta, (\theta * \phi)\} \text{ where } \theta, \phi \in \mathcal{ESL}_n, i \in N, \\
&\quad * \in \{\wedge, \vee, \rightarrow\} \\
\mathcal{ESL} &= \bigcup_{n \in \mathbb{N}} \mathcal{ESL}_n.
\end{aligned}$$

Sentence $K_i(\theta)$ reads "agent i knows θ", (i.e. knows that θ is true). Note that $K(\cdot)$ is *not* a compositional connective, in the sense outlined in Section 2 above (exercise!). We say that *agent i knows θ at state s* if $K_i(\theta)$ is satisfied at s.

Definition 4.1 (Epistemic satisfiability). State $s \in S$ satisfies $K_i(\theta)$, written $s \models K_i(\theta)$ if and only if $v(\theta) = 1$ in all states $t \in S$ such that $R(s, t)$.

It is immediate to note that this definition *extends* that of classical consequence, hence the abuse of notation. The idea is that classical consequence holds "locally" at each state, whereas epistemic validity requires a "global" check on all accessible states. As anticipated above, distinct constraints on accessibility give rise to distinct formalisations of knowledge. The standard epistemic logic $S5$ arises by imposing that $R \subseteq 2^S$ be an equivalence relation.

So far we have considered only one of the two standard presentations of logical consequence, namely the one based on the (preservation of) satisfiability which constitutes the backbone of Definition 3.1 and Definition 4.1. Those semantic (as logicians say) notions of consequence are coupled with *axiomatic* counterparts. It is this coupling which enables the AxAP perspective on the logical investigations on the concept of knowledge.

The axioms of $S5$ include all propositional tautologies, in addition to the following:

$$\begin{aligned}
\mathbf{K}: &\quad K_i(\theta \to \phi) \mid (K_i\theta \to K_i\phi) \\
\mathbf{T}: &\quad K_i\theta \mid \theta \\
\mathbf{4}: &\quad K_i\theta \mid K_iK_i\theta \\
\mathbf{5}: &\quad \neg K_i\theta \mid K_i\neg K_i\theta
\end{aligned}$$

where the stroke symbol is to be read intuitively as "entails". So, the above mentioned axiom **T** reads as: *if agent i knows θ then θ is true.* Axiom **4**, also mentioned above, prescribes that agents should know that they know what they know, i.e. should be able to *positively introspect.* The next axiom demands even more, that if there's a θ they don't know then they should know that they don't know θ.

The definition of epistemic satisfiability (i.e. what it means for you to know θ at state s) is given independently of the actual properties which are satisfied by R. Different constraints on R lead to materially different formalisations of knowledge. But in virtue of correspondence, this means that different accessibility relations are captured by different axioms, i.e. they capture properties of knowledge which may or may not be deemed adequate for a normative characterisation of knowledge. So you can see how epistemic logic provides an incredibly rich and flexible framework for the investigation of what we should take knowledge to be.

One immediate advantage of the axiomatic presentation of epistemic logic(s) is that it allows for distinguishing among epistemic attitudes. Popular ones are

(1) i knows θ
(2) i believes θ
(3) i is certain that θ
(4) i is aware of θ

(5) i is informed that θ.

The question as to which are the fundamentally independent attitudes feeds much of the current research in epistemic logics. Of particular interest to economic theory are the following. First, whether knowledge entails belief. Second, whether awareness can be used to model the epistemic attitudes of *non omniscient* agents.

To conclude, in analogy with Section 3, suppose you are dissatisfied with the consequences of the "common knowledge" assumption in the definition of Nash equilibrium. One reaction is to consider "correlated equilibrium" as being less demanding, as it only requires rational players to have common beliefs in the form of common prior probabilities. Epistemic logics provide an interesting framework to investigate *qualitative* analogues of the notion. AxAP suggests to undertake a more radical step and to question –at the very root– our conception of knowledge.

4.2. **Further reading.** (Fagin et al., 1996) is the first systematic formalization of the epistemic interaction of logical agents. (Shoham and Leyton-Brown, 2009) builds on its tradition.

Readers with a background in mathematical logic are referred to (Blackburn et al., 2001) for a state-of-the-art introduction to Modal logic and (Blackburn et al., 2007) for a comprehensive account of its development. (Meyer and van der Hoek, 1995) shows how by the mid-1990s epistemic logic was a standard analytical tool in artificial intelligence.

On the logic of "being informed" see Allo (2011). Halpern and Pucella (2007) provides an overview of the problem of logical omniscience. Schipper (2012) is a recent bibliography on (un)awareness.

References

E.W. Adams. *A Primer of Probability Logic.* CSLI Lecture Notes, 1996.

P. Allo. The logic of 'being informed' revisited and revised. *Philosophical Studies*, 153(3):417–434, February 2011.

R.J. Aumann. Agreeing to disagree. *Annals of Statistics*, 4:1236–1239, 1976.

P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic.* Cambridge University Press, June 2001.

P. Blackburn, J. Van Benthem, and F. Wolter. *Handbook of modal logic.* Elsevier, 2007.

A. Bochman. *A Logical Theory of Non Monotonic Inference and Belief Change.* 2001.

G Boolos, J.P. Burgess, and R.C. Jeffrey. *Computability and Logic*. Cambridge University Press, 4th edition, 2002.

I. Chiswell and W. Hodges. *Mathematical Logic*. Oxford University Press, 2007.

F. Ciraulo, M. E. Maietti, and P. Toto. Constructive version of Boolean algebra. *Logic Journal of IGPL*, 21(1):44–62, 2013. ISSN 1367-0751.

F. Dietrich and C. List. Arrow's theorem in judgment aggregation. *Social Choice and Welfare*, 29(1):19–33, October 2007. ISSN 0176-1714.

H. B. Enderton. *A Mathematical Introduction to Logic*. Academic Press, 2nd edition, 2001.

R. Fagin, J.Y. Halpern, M. Yoram, and M. Vardi. *Reasoning about Knowledge*. MIT Press, 1996.

M. Fedel, H. Hosni, and F. Montagna. A logical characterization of coherence for imprecise probabilities. *International Journal of Approximate Reasoning*, 52(8):1147–1170, 2011.

R. Haenni, J.W. Romeijn, J. Williamson, and G. Wheeler. *Probabilistic Logics and Probabilistic Networks*. Springer, 2011.

P.R. Halmos and S.R. Givant. *Logic as algebra*. The Mathematical Association of America, 1998.

J.Y. Halpern and R. Pucella. Dealing with logical omniscience. In *Proceedings of the 11th conference on Theoretical Aspects of Rationality and Knowledge*. ACM, 2007.

W. Hodges. *Logic*. Penguin Books, 1985.

R.C. Jeffrey. *Subjective probability: The real thing*. Cambridge University Press, 2004.

L. A. Kornhauser and L. G. Sager. Unpacking the court. *Yale Law Journal*, 96(1):82–117, 1986.

C. List. The theory of judgment aggregation: an introductory review. *Synthese*, 187(1):179–207, December 2011. ISSN 0039-7857.

J.-J.Ch Meyer and W. van der Hoek. *Epistemic Logic for AI and Computer Science*. Cambridge University Press, 1995.

J.B. Paris. *The uncertain reasoner's companion: A mathematical perspective*. Cambridge University Press, 1994.

J.B. Paris. A Note on the Dutch Book Method. In Teddy Seidenfeld (Eds.): Gert De Cooman, Terrence Fine, editor, *ISIPTA '01, Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications, Ithaca, NY, USA*. Shaler, 2001.

J.B. Paris and A. Vencovska. *Pure Inductive Logic*. Cambridge University Press, 2013.

P. Pettit. Deliberative democracy and the discursive dilemma. *Philosophical Issues*, 11:268–299, 2001.

B. Schipper. Unawareness bibliography. *http://www.econ.ucdavis.edu/faculty/schipper/unaw.htm*, 2012.

Y. Shoham and Kevin Leyton-Brown. *Multiagent systems: algorithmic, game-theoretic, and logical foundations.* Cambridge University Press, 2009.

J. Williamson. *In defence of objective Bayesianism.* Oxford University Press, 2010.

Scuola Normale Superiore, Pisa and Centre for Philosophy of Natural and Social Science, London School of Economics, London

E-mail address: h.hosni@gmail.com