

# Common Knowledge and Convention

Giacomo Sillari

Published online: 1 July 2008  
© Springer Science+Business Media B.V. 2008

**Abstract** This paper investigates the epistemic assumptions that David Lewis makes in his account of social conventions. In particular, I focus on the assumption that the agents have common knowledge of the convention to which they are parties. While evolutionary analyses show that the common knowledge assumption is unnecessary in certain classes of games, Lewis' original account (and, more recently, Cubitt and Sugden's reconstruction) stresses the importance of including it in the definition of convention. I discuss arguments pro et contra to argue that, although the assumption might be relevant to a descriptively adequate account of social convention, it is not required for its rational reconstruction. I then point out that Lewis' account, properly speaking, is of common reason to believe, rather than of common knowledge, and argue that, in order to formalize aptly the distinction between reason to believe and belief, standard formal epistemic models need to be supplemented with so-called awareness structures. Finally, I stress that the notion of knowledge implicit in Lewis' text involves interesting elements that cannot be captured in the standard propositional formalizations, but need the full expressive force of quantified epistemic logic.

**Keywords** David Lewis · Convention · Common knowledge · Awareness · Quantified epistemic logic

## 1 Lewis on Convention

The idea that games of coordination lie at the core of the notion of convention is a crucial insight in Lewis' account

of convention. It stems from Thomas Schelling's *The Strategy of Conflict* (Schelling 1960). In *Convention*, David Lewis elaborates on the insight. One of the most fecund (yet controversial) developments put forth in (Lewis 1969) consists of the introduction and analysis of epistemic elements in conventional situations, that is the study of the relation between (common) knowledge and social conventions.

In coordination games players' interests (roughly) coincide while in zero-sum games the interest of the players are always opposed.<sup>1</sup> Thus, coordination problems (non-trivial, in that they possess more than one strict Nash equilibrium) underlie every convention, since a convention is one particular recurrent equilibrium of such games.

When addressing the question of how a specific equilibrium can be reached, Lewis introduces the idea of a *system of mutual expectations*. In general, in order to have a sufficient reason for choosing a particular action, an agent needs to have a belief (up to a sufficient degree) that the other agent will choose a certain action. Lewis argues that in the case of coordination games such a sufficient degree of belief is justified by a system of mutual expectations. The focus of Lewis' study is on how conventions are sustained, rather than on how they originate. Although, according to Lewis, in both cases coordination is achieved by means of a system of mutual expectation, the reconstruction of the origin of conventions, and the justification of their maintenance differ in the process leading to the creation of the expectations. A system of mutual expectations can come about in different ways. A natural one is,

---

G. Sillari (✉)  
Philosophy, Politics and Economics Program,  
University of Pennsylvania, Philadelphia 19104, USA  
e-mail: gsillari@sas.upenn.edu

<sup>1</sup> The distinction was pointed out a few years before *Convention* by Schelling, cf. Schelling (1960), p. 84: "If the zero-sum game is the limiting case of pure conflict, what is the other extreme? It must be the "pure collaboration" game in which players win or lose together, having identical preferences regarding the outcome."

for example, that of agreeing on playing a certain strategy profile. Another common coordination device is salience. Facing a new coordination problem, the agents recognize that an equilibrium has certain salient features, and each player expects them to be noticed by the other players too. A particular kind of salience is precedent: in this case the salient trait of the equilibrium is that it has served as a solution of a similar coordination problem in the past. Although Lewis does not state the point explicitly, it seems reasonable to interpret his text as saying that salience, in general, may serve as a coordination device for originating conventions, whereas precedent is the coordination device involved in their perpetuation.<sup>2</sup>

Since systems of mutual expectations play such a fundamental role in Lewis' theory of convention, it is natural to investigate what mechanisms produce them. Lewis elucidates this point by providing a definition of "common knowledge." Common knowledge is that epistemic state in which all agents know that  $p$ , all agents know that all agents know that  $p$ , all agents know that all agents know that all agents know that  $p$ , and so on *ad infinitum*. The notion of common knowledge plays important roles in a variety of fields.<sup>3</sup>

The following is Lewis' definition:

**Definition 1.1** A proposition  $p$  is common knowledge in the group  $G$  if a state of affairs  $A$  obtains such that

1. Everyone in  $G$  has reason to believe that  $A$  holds;
2.  $A$  indicates to everyone in  $G$  that everyone in  $G$  has reason to believe that  $A$  holds;
3.  $A$  indicates to everyone in  $G$  that  $p$ .

Lewis' text<sup>4</sup> suggests that the relation of indication is not material implication, incorporating some kind of inductive inference. In general, if  $A$  indicates  $x$  to  $i$ , this means that, if  $i$  had reason to believe that  $A$  holds,  $i$  would thereby have reason to believe that  $x$  holds as well. Clauses (1)–(3), along with suitable assumptions about the agents' reasoning capabilities and inductive standards, originate an infinite series of epistemic propositions such that everyone in  $G$  has reason to believe that  $p$ , everyone has reason to believe that everyone has reason to believe that  $p$ , and so

<sup>2</sup> Most authors endorse this interpretation, cf. e.g. Cubitt and Sugden (2003) or Skyrms (2003).

<sup>3</sup> It is difficult to overestimate the influence that the introduction of such an idea has exerted in so many different fields, ranging from economics (Geanakoplos 1992) to computer science (Fagin et al. 1995; Meyer and van der Hoek 1995), from logic (besides the propositional results of Fagin et al. 1995 and Meyer and van der Hoek 1995, cf. also issues of quantification in Wolter 1999 and Sturm et al. 2002, and the proof theoretical analysis of Alberucci and Jaeger 2004) to linguistics (Clark 1996).

<sup>4</sup> For a more detailed analysis of this claim, cf. both Cubitt and Sugden (2003) and Sillari (2005).

on. The state of affairs  $A$ , which allows the agents in the group  $G$  to have common knowledge of the proposition  $p$ , is said to be a *basis* for common knowledge of  $p$  in  $G$ .

Lewis' definition of convention then is the following (cf. Lewis 1969, p. 58):

**Definition 1.2** A regularity  $R$  in the behaviour of members of a population  $P$  when they are agents in a recurrent situation  $S$  is a convention if and only if it is true that, and it is common knowledge in  $P$  that, in any instance of  $S$  among members of  $P$ ,

- (1) everyone conforms to  $R$ ;
- (2) everyone expects everyone else to conform to  $R$ ;
- (3) everyone prefers to conform to  $R$  on condition that the others do, since  $S$  is a coordination problem and uniform conformity to  $R$  is a coordination equilibrium in  $S$ .

The definition is further refined in subsequent chapters of Lewis' study, but, for the purposes of this essay, it is not necessary to consider such refinements here.

## 2 Convention Without Common Knowledge

The requirement in the above definition that its clauses be common knowledge in the population is perplexing for several reasons. First and foremost, from a purely game-theoretical point of view, there is no need to require common knowledge of the convention in a population  $P$ . All that matters for the agents in  $P$  is that they expect all other agents to stick to their part in a specific equilibrium profile. Once this first order expectation is available, rational players will coordinate on the specific equilibrium profile. In fact, it is objected to Lewis that common knowledge of conformity does *not* justify equilibrium play, and it is doubtful (cf. Bicchieri 1993, p. 66 ff.) that common knowledge of the structure of the game and of the agents' game-theoretical rationality is sufficient for obtaining coordination. Also, in the evolutionary game-theoretic analysis of certain classes of coordination games, the common knowledge requirement is not necessary to achieve (or sustain) coordination in the population, and hence can be dispensed with (cf. Skyrms 1996, 2003). Second, even if we admit that common knowledge is sufficient to attain coordination, it is not entirely clear how the agents might achieve common knowledge. To be sure, definition 1.1 above provides us with a finite set of conditions from which the infinite hierarchy of epistemic states constituting the agents' common knowledge can be inferred. But what I referred above to as "suitable assumptions about the agents' reasoning capabilities" (that is to say, what Lewis refers to as "ancillary premises", cf. Lewis 1969, p. 55, about the agents' inductive standards,

background information and rationality<sup>5</sup>) are far from innocuous. They entail that agents are symmetric reasoners,<sup>6</sup> so that if a state of affairs indicates *p* to someone, it indicates *p* to all other agents under consideration, and the agents know this. Thus, in the definition of common knowledge itself there seems to be a problematic requirement of some sort of pre-existing common knowledge. Third, even if we admit that such ancillary premises are acceptable, the question of how can we sensibly ascribe an infinite number of knowledge states to non-ideal agents still stands. Lastly, even if the infinite regress implicit in Lewis' definition can be made harmless, it is not immediately clear whether the definition applies to a degree of generality sufficient to explicate the pervasive phenomenon of social norms and conventions.

Several authors, albeit sharing with Lewis the view that coordination is at the core of the phenomenon of social conventions, and agreeing with him that the game-theoretical explanation is the most suitable to make sense of the phenomenon, reject the idea that conventions need to be common knowledge in the population in which they hold.

For example<sup>7</sup> analyzing the special case of conventions that deals with *signaling games*, Brian Skyrms observes in (Skyrms 1996) that, without assuming anything at all about the agent's rationality and common knowledge, but rather having the individuals "programmed" to play specific strategies and the overall evolution of the population directed by the replicator dynamics, agents never fail to coordinate on *some* signaling equilibrium<sup>8</sup> of the game. Which signaling equilibrium is selected depends on the initial conditions of the system, or on the presence of noise or other kinds of random fluctuations, not on the agents' cognitive or epistemic states. The advantage of an evolutionary analysis of signaling games and of coordination games in general, are various. For one thing, the evolutionary analysis does not presuppose any higher cognitive ability of the agents, and therefore it applies to humans as well as to the rest of the biological world. By dispensing with any kind of shared knowledge, it avoids some

of the pitfalls associated with Lewis' analysis (in particular, it avoids issues of circularity in the definition of common knowledge in which, cf. above, pre-existing common standards of reasoning are assumed), whereas, by dispensing with cognitive notions such as that of salience, the evolutionary analysis applies also to those cases that, being perfectly symmetric, present the agents with no action standing conspicuously out. Moreover, if we do presuppose that certain actions are in fact salient, the evolutionary analysis itself could prove useful for the determination of the evolution of the psychological propensities that make us consider the characteristics of certain actions as salient. In a sense, thus, Skyrms' analysis provides us with a naturalization of the concept of salience. However, the evolutionary analysis does not exclude other kinds of analyses in which epistemic and cognitive assumptions about the agents do play a role (cf. Skyrms 2003, p. 56) and could actually be augmented or complemented by those. In sum, although assumptions involving common knowledge and salience are not necessary to model the emergence and maintenance of coordinating behavior in evolutionary models, they can serve the purpose of informing more realistic evolutionary models, for example by determining constraints on the initial population or on the dynamics of the system. Also, they can function as key notions in analyses focused on the agents' cognitive capability. Cubitt and Sugden argue in (Cubitt and Sugden 2003) that Lewis' original project was of this kind, and maintain that such an analysis could prove fruitful.

The close relation between social conventions and social norms is already stressed in Lewis' essay<sup>9</sup>, and it is further explored in Ullman-Margalit (1977) and Sugden (1986). More recently, Cristina Bicchieri (2005) has provided an analytical study of social norms in the tradition pioneered by Lewis. Although her analysis contains various elements of evolutionary theory, it moves from a static definition of social norms, much in the spirit of *Convention*. In light of some of the observations that I mentioned at the onset of this section, her definition of social norm does not contain a common knowledge clause. In particular Bicchieri, in her *Grammar of Society*, is keen to stress that the common knowledge assumption in Lewis' definition of convention does *not* provide the agents with a reason to conform to the coordination equilibrium, but rather provides an epistemic justification of the agents' reason to conform.<sup>10</sup> However,

<sup>5</sup> The notion of rationality Lewis is referring to here is not the game-theoretical one, but rather, cf. Bicchieri (1993), *epistemic* rationality. On this subject, cf. also Sillari (2005).

<sup>6</sup> Cf. Vanderschraaf (1998).

<sup>7</sup> The examples given here do not mean to be exhaustive or especially representative of the contributions given by game theorists to the study of convention or to the formalization of epistemic notions. I believe however that they are representative of the game theoretic literature that rejects the notion that conventions need to be common knowledge (other apt examples could have been Ken Binmore, or Peyton Young), as opposed to those authors (for instance, Bob Sugden, or Herb Gintis) who see common knowledge as an essential constitutive trait of social conventions.

<sup>8</sup> Signaling games are a special kind of coordination games in which various equilibria exists. In some of the equilibria, the agents achieve effective communication. Such equilibria are called signaling equilibria. Cf. Lewis (1969), Crawford and Sobel (1982).

<sup>9</sup> Cf. Lewis (1969, p. 97): "conventions may be a species of norms: regularities to which we believe one ought to conform. I shall argue that they are."

<sup>10</sup> More details of this argument are given in the next section, and in Sillari (2005). Bicchieri argues more extensively against the introduction of a common knowledge clause in the definition, in Bicchieri (2005, pp. 36-37), supporting the conclusion that in a *rational reconstruction* of the notion of social norm (or convention) the common knowledge assumption is not necessary.

in her account notions like cognitive focus, psychological cues or norm priming— notions that seem to presuppose assumptions about what agents know about the situation and about each other—are essential components to the explication of normative behavior. Such elements are reconciled with the absence of epistemic clauses in the definition of social norms by the observation that norm abidance, oftentimes, is automatic rather than conscious and deliberate. Thus, although unnecessary in a general rational reconstruction of the notion of social norm, common knowledge and salience are contingencies that possibly accompany instances of normative or conventional behavior.

### 3 Common Knowledge and Convention

Let us now turn back to David Lewis' original study, to try and understand why he included common knowledge as a necessary element of his definition of convention. In *Convention*, Lewis mentions two reasons why common knowledge should be included in the definition of convention. The first relies on purely descriptive grounds, since common knowledge is a relevant characteristic of conventions<sup>11</sup> and Lewis wants to incorporate in his rational reconstruction of convention all relevant facts pertaining to it. The second reason why Lewis incorporates common knowledge into his definition of convention is that it prevents certain odd situations from counting as conventions as they would in the absence of the common knowledge requirement. Suppose, for example, that everyone drives on the right to avoid collisions, but at the same time everyone thinks everyone else drives on the right for frivolous, entirely unrelated reasons. If the clauses of the definition of convention are not known by the agents, this situation, which intuitively we want excluded from the number of rational reconstruction conventional situations, fulfills the definition of convention. But, says Lewis, it is not sufficient that we amend the definition by requiring that each agent knows everyone else's preferences for avoiding collisions. This is because the counterexample just set forth can be reiterated at a higher level of knowledge: Suppose everyone drives on the right to avoid collisions and everyone thinks everyone else drives to the right for that reason. However, suppose that now everyone thinks that everyone thinks that everyone else drives to the right for frivolous, entirely unrelated reasons. Hence, if the clauses of the definition are not known to be known by the agents, this situation counts,

countering our intuition, as a convention. Clearly the construction can be replicated for arbitrary high levels of agents' knowledge, and the definition becomes impervious to counterexamples of this kind only if its clauses are assumed to be common knowledge among the agents.<sup>12</sup>

Of course, an important reason why Lewis sees common knowledge as such a relevant characteristic of social conventions lies in the fact that common knowledge (or, as he also refers to this notion in Lewis 1969, a system of mutual expectations) functions as a *possible* justification grounding the agents' first-order expectations about each other's behavior in the coordination game. Skeptics of Lewis' account like Margaret Gilbert aim at this aspect (among others) when critiquing his game-theoretical analysis of convention. The gist of the argument is that each level of mutual knowledge is justified by the level immediately above it and therefore, since in the case of common knowledge the hierarchy of levels is infinite, common knowledge cannot be invoked to conclusively justify any first-order belief (and, *a fortiori*, no course of action can justifiably be taken). This may be clearer if illustrated by a simple example: If  $x$  represents doing one's part in a coordination equilibrium, Ann is justified in doing  $x$  because she thinks that Bob will do  $x$ . But Ann is justified in thinking that Bob will do  $x$  by thinking that Bob thinks that Ann will do  $x$ . But Ann is justified in entertaining this last belief if she thinks that Bob thinks that Ann thinks that Bob will do  $x$ , and so on. Since the ascending chain is infinite, no conclusive grounds can be given for Ann's first order belief that Bob will do  $x$ . Of course, an argument mirroring the one just given holds for Bob, too. Hence, common knowledge cannot provide game-theoretically rational agents with a conclusive reason to act. However, the process of justifying agents' *beliefs* does not pertain to the sphere of game-theoretical rationality, but rather to that of their epistemic rationality—to the process meant to justify the agents' beliefs, not to the process meant to justify their actions.<sup>13</sup> But as an instrument of theoretical, rather than practical reason, the common knowledge

<sup>11</sup> Cf. Lewis (1969, p. 59): "common knowledge of the relevant facts seems to be one [important feature common to our examples of conventions]." For illuminating examples of common knowledge and of its relevance in our social practices and in general for anthropological syntheses, cf. Chwe (2001).

<sup>12</sup> This kind of "regress argument" for the introduction of common knowledge is not unique to the Lewisian definition of social convention. For example Schiffer (1972) justifies his emendation to Grice's definition of non-natural meaning (1989) through a similar argument. In a nutshell: suppose that the recognition of the speaker's intention in an act of communication is essential for successful communication. Then the hearer knowing that the speaker, in uttering  $p$ , intended  $x$  is necessary to establish meaning of  $p$ . But the speaker must also have intended that the hearer recognize that the speaker intended  $x$ , hence etc. The introduction of a common knowledge clause eliminates the vicious regress. Clark and Marshall (1981) justify the claim that speaker and audience have to have common knowledge of the reference of definite descriptions through a similar regress argument.

<sup>13</sup> This argument is detailed in Sillari (2005).

assumption gives to the agents, in certain occasions, reason to believe that the other party will do  $x$ . This belief, in turn, is necessary to motivate coordinative action.<sup>14</sup>

The relation between attainment of common knowledge and coordinative action is explored in Cubitt and Sugden's article on *Convention* (cf. Cubitt and Sugden 2003). In it, they suggest that Lewis provides in the essay "an astonishingly comprehensive, consistent and original theory of practical reason in the context of recurrent games. Lewis' approach offers an alternative both to the modeling strategy of classical game theory, in which self-contained games are played by hyper-rational agents, and to that of evolutionary game theory, in which players' behaviour is the product of blind processes of selection." (cf. Cubitt and Sugden 2003, p. 203.) In Cubitt and Sugden's reconstruction, common knowledge is the key notion for appreciating the originality of Lewis' approach. On the one hand, endowing agents with common knowledge (intended in the sense of Lewis) requires equipping them with reasoning capabilities beyond those of the "hyper-rational agents" of traditional game theory, for common knowledge is entwined with the notion of salience and the latter cannot be understood in purely game-theoretical terms. On the other, the agents' behavior cannot simply be "the product of blind processes of selection", since the cognitive abilities of the agents do play such a crucial role in the reconstruction. To better appreciate and evaluate Cubitt and Sugden's argument, however, we now take a closer and more formal look at the definition of common knowledge.

#### 4 On the Definition of Common Knowledge

If the epistemic status of the agents and, in particular, the epistemic status to which the expression 'common knowledge' refers to is, in certain cases, part of the definition of convention, the questions of both its characterization and its origin arise. I argue in this section that the answers to such questions have been given in the literature following the publication of *Convention* are already *in nuce* in Lewis' text.

Echoing the analysis of the interrelations between different approaches to common knowledge provided by Jon Barwise (1988), I shall refer to the *iterative* and *fix point* accounts of common knowledge.<sup>15</sup>

As mentioned in the opening section, common knowledge is the special kind of knowledge enjoyed by a group of agents  $G$  when all agents in  $G$  know that  $p$ , all agents in  $G$  know that all agents in  $G$  know that  $p$ , and so on *ad infinitum*. This *iterative* characterization of common knowledge is intuitive, but at the same time quite informal. More importantly, consisting of an infinite list of epistemic clauses, and ascribing to the agents an infinite amount of epistemic states, it is clearly inadequate from the descriptive standpoint. In other words, although the iterative definition of common knowledge makes it intuitively clear what common knowledge *is*, it does not explain how common knowledge comes about; nor does it do a good job at describing the actual epistemic state of the agents in the group under consideration.

In the philosophical literature, Harman is generally credited with primacy regarding the *fix point* characterization of common knowledge. His characterization appears in a review (cf. Harman 1977) of Bennett's *Linguistic Behavior*, and indeed the formulation serves the purpose of amending Grice's definition of non-natural meaning by avoiding the "regress argument" mentioned above, as it does in Schiffer (1972). Thus, Harman defines common knowledge of  $p$  in a group as the situation in which "each knows  $p$  AND WE KNOW THIS, where the THIS refers to the whole fact known." One could say that common knowledge arises in situations in which there is a *public* statement, where a statement is said to be public<sup>16</sup> when all agents can hear and understand the statement and they know this. To see this more clearly, it is useful to introduce some formalism. Let  $K_i$ ,  $E_G$  and  $C_G$  denote, respectively, "agent  $i$  knows", "all agents in  $G$  know" and "it is common knowledge in  $G$  that", while  $\varphi$ ,  $\psi$ , ... denote propositional metavariables and  $\wedge$ ,  $\neg$ ,  $\rightarrow$  are the logical symbols for conjunction, negation and material implication, respectively. A *normal* epistemic system contains the axioms in the list on the left hand side, along with some combination of the remaining four:

PC	All axioms of propositional calculus	T	$K_i \rightarrow \varphi$
K	$(K_i \varphi \wedge K_i(\varphi \rightarrow \psi)) \rightarrow K_i \psi$	4	$K_i \varphi \rightarrow K_i K_i \varphi$
MP	From $\varphi$ and $\varphi \rightarrow \psi$ , infer $\psi$	5	$\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$
NEC	From $\varphi$ , infer $K_i \varphi$	D	$\neg K_i(\varphi \wedge \neg \varphi)$

<sup>14</sup> Whether it is in fact the case that common knowledge functions as a theoretical reason to believe in the coordinative performance of the other party remains an empirical question, and possibly an interesting subject of empirical scrutiny.

<sup>15</sup> The distinction is analyzed in greater detail in (Vanderschraaf and Sillari 2005).

<sup>16</sup> The notion of a public event was probably introduced by Paul Milgrom in (Milgrom 1981). There it is shown that common knowledge can be axiomatically determined, one of the axioms requiring in fact that public events be common knowledge.



Systems meant to represent *knowledge* usually include all of the axioms above excluding D, whereas systems meant to represent belief exclude T.<sup>17</sup> In order to represent group knowledge, the following definitory axioms are added to the system.

$$\begin{array}{l} \text{E} \quad E_G \varphi \leftrightarrow \bigwedge_{i \in G} K_i \varphi \\ \text{C} \quad C_G \varphi \leftrightarrow E_G(\varphi \wedge E_G \varphi) \end{array}$$

Note that since the language of epistemic logic is finitary, common knowledge cannot be expressed as the infinite conjunction of formulas  $E_G \varphi \wedge E_G E_G \varphi \wedge \dots$ , as it can instead be done in the economics literature<sup>18</sup>. Finally, a rule to *derive* common knowledge is needed. The following *induction rule* is used:

I From  $\varphi \rightarrow E_G(\varphi \wedge E_G \psi)$ , infer  $\varphi \rightarrow C_G \psi$

The induction rule tells us (in agreement with the intuition about what a “public situation” is) that if the public situation  $E_G(\varphi \wedge E_G \psi)$  follows from  $\varphi$ , then one can derive that common knowledge of  $\psi$  also follows from  $\varphi$ . It is important to notice that this rule, as well as all the axioms above, are mere syntactic symbols. They can be given a semantics in terms of so-called Kripke structures or, equivalently, in set-theoretical terms.<sup>19</sup>

It is interesting to see that the fix point approach is equivalent to the iterative approach in the case of *normal* systems of epistemic logic (those systems assuming distributivity of the knowledge operator over implication, i.e. axiom K above), while it is stronger than the iterative approach in systems weaker than normal ones.<sup>20</sup> Besides the mathematical significance of this fact, it also seems to lend support to the idea that, as much as the iterative characterization remains a good *description* of the phenomenon of common knowledge, the appropriate *definition* of common knowledge has to be given in terms of public events, as it is the case in the fix point characterization.

The vast majority of authors who contributed to the early discussion on common knowledge share the view that the iterative definition is not adequate to represent formally

the notion of common knowledge, although at the same time they all agree that it does capture the informal intuition behind it.<sup>21</sup> Although none of these authors resort directly to the fix point definition of common knowledge, their accounts all share the same intuition that underlies the fix point definition, i.e. that common knowledge is best understood as proceeding from a public event. Consider for example Lewis’ definition of common knowledge (see definition 1.1 above). If, to simplify the discussion, we reduce Lewis’ indication to material implication, the definition contains the two clauses

- (i)  $A \rightarrow E_G A$ ,
- (ii)  $A \rightarrow p$

as the two clauses that *generate* common knowledge of  $p$  in  $G$  from  $A$ . Obvious propositional reasoning compacts the two clauses in  $A \rightarrow (E_G A \wedge p)$ . Hence the aforementioned *induction rule* of epistemic logic (from  $\varphi \rightarrow (E_G \varphi \wedge \psi)$ , infer  $\varphi \rightarrow C_G \psi$ ) is implicit in Lewis’ account of common knowledge.

Stephen Schiffer (1972, pp. 34–35) gives a characterization of common knowledge that is fairly similar to Lewis’. It can be rephrased by the following two clauses:

- (i-s) in the current situation, everyone in  $G$  is normal, and being normal is sufficient to know that everyone who is normal is in fact normal;
- (ii-s) in the current situation, everyone normal has sufficient reason to know that they have sufficient reason to know that  $p$ .

Normality guarantees, in clause (i-s), that the inference from the current situation to knowledge that  $p$  is warranted, while clause (ii-s) guarantees that conditions sufficient to generate iterated knowledge of arbitrary level are satisfied.<sup>22</sup> It is straightforward to see that, if we denote “situation sufficient for the agents in  $G$  to infer that they

<sup>17</sup> Thus, systems representing knowledge include axiom T, which states that anything known by an agent is actually the case, while systems representing belief allow agents to entertain false belief, while requiring, by axiom D, that the set of belief they entertain be consistent. Axiom D is not required in the system representing knowledge because it is easily derived from T.

<sup>18</sup> Some authors do consider infinitary epistemic systems: cf. e.g. Heifetz (1999), Kaneko and Nagashima (1996), Alberucci and Jaeger (2004).

<sup>19</sup> The equivalence between the two approaches carries over to the semantics of the group knowledge operators, implying that Aumann’s formal definition of common knowledge is equivalent to the fix point characterization.

<sup>20</sup> For a formal proof, cf. Heifetz (1999). The observation was already made, without proof, in Gilbert (1989, p. 468, n. 30.).

<sup>21</sup> Lewis (1969, p. 52) asks “[w]hat premises have we to justify us in concluding that others have certain expectations, that others expect others to have certain expectations, and so on?”; Schiffer (1972, pp. 32 ff.), after introducing common knowledge, which he calls *mutual knowledge\**, through the iterative definition, gives a finite set of “conditions which must obtain for [common knowledge] to be realized”; Aumann (1976, p. 1237) shows that “the formal definition of ‘common knowledge’ is equivalent to the informal description; Clark and Marshall (1981) also introduce the notion through its iterative description, but then devote a substantial part of their article to dissect a number of variations on the coming about of common knowledge: such variations can in general be subsumed in the Lewisian approach.

<sup>22</sup> To appreciate the similarity with Lewis’ account, consider that for Lewis the inference from the current situation to knowledge that  $p$  is warranted by the relation of indication *and* by the fact that agents share inductive standards and background information. Thus, the assumption of common inductive standards corresponds to Schiffer’s “normality” condition and indication to Schiffer’s “sufficiency for iterated knowing.”

are normal” with *A*, then clauses (i-s) and (ii-s) are rendered formally by clauses (i) and (ii) above.

A further characterization of common knowledge can be found in Clark and Marshall’s article on definite reference (cf. Clark and Marshall 1981). In the essay, meant to show that common knowledge plays a decisive role in communicating and understanding expressions containing definite descriptions, the authors discuss at length several varieties of common knowledge (lasting as opposed to temporary knowledge, knowledge of particulars as opposed to knowledge of generalizations, etc.). They provide a rather general characterization of common knowledge in terms of *community membership*.<sup>23</sup> There are various degrees and kinds of community membership. For example, in a movie theatre the public at a certain show would constitute a community. Or, more generally, speaker of the same language, or educated citizens of the same country also constitute a community in Clark and Marshall’s sense. In such communities, various items of knowledge are common knowledge among their members and, in a sense not distant from Lewis’, membership in the community consists of holding certain items of common knowledge.<sup>24</sup> In particular, Clark and Marshall observe, “[common] knowledge of community membership makes an excellent basis *G* for the [common] knowledge induction schema.” Two assumptions are needed for community membership to originate common knowledge<sup>25</sup>:

- (i-cm) being a member in the community indicates that everyone else in the community is a member of it
- (ii-cm) being a member in the community gives one sufficient reason to know *p*.

Again, modulo the inferential process which (as in the cases of Lewis’ and Schiffer’s) is not confined to material implication, the two clauses can be rendered as (i) and (ii) above, revealing a substantial structural similarity among these accounts of common knowledge.

These characterizations, thus, enjoy two relevant aspects. On the one hand, they tend to have more or less pronounced realistic features meant to capture the actual process of common knowledge acquisition in a group of agents. On the other, they can be understood, at a more idealized level, as instances of the mathematical

approaches to representing knowledge adopted by logicians and economists. Cubitt and Sugden deny this latter characteristic of Lewis’ common knowledge, claiming that the epistemic notions recoverable in *Convention* cannot be satisfactorily formalized through the standard mathematical models.<sup>26</sup> Thus, in Cubitt and Sugden (2003) they conduct their reconstruction and evaluation of Lewis’ theory of convention, and of the role that the common knowledge assumption plays in it, along the lines of the former characteristic. They strive to make formal sense of the process through which the agents, in Lewis’ account, acquire common knowledge and, in so doing, explicate how conventions perpetuate themselves. In their analysis, common knowledge, precedent and inductive reasoning work closely together to justify conventional behavior. The idea<sup>27</sup> is that the state of affairs *A* (which exhibits past conformity to a certain coordination equilibrium) functions as a *basis* from which, possibly through inductive reasoning, common knowledge of *p* (the event of conformity to the previously successful coordination equilibrium) is inferred. The precedent exhibited by *A*, in order to be understood as such by the agents involved in the coordination problem, needs to be based on a projectible regularity, allowing the agents to draw the appropriate inferences that justify coordinating behavior.

## 5 Knowledge and Awareness

Although the term “common knowledge” has a remarkable fortune in the literature of so many academic fields, it is noteworthy that Lewis is interested in expectations, that is to say *beliefs*, rather than knowledge. Moreover, his definition does not even immediately deal with beliefs, but rather with reasons to believe, being in fact a definition of “common reason to believe” a certain proposition. However, the term ‘common knowledge’ became prevalent, perhaps also because the first (and seminal) mathematical formulation of the concept<sup>28</sup> had a natural interpretation in terms of knowledge rather than belief. Lewis himself later acknowledges the incongruence: “That term [common knowledge] was unfortunate, since there is no assurance that it will be knowledge, or even that it will be true” (Lewis 1978, p. 44, n. 13). I take it that here Lewis is pointing out that reasons to believe may end up becoming *false* beliefs about the world (“there is no assurance that it

<sup>23</sup> This idea is elaborated in much greater detail in Clark (1996, pp. 92–125).

<sup>24</sup> Cf. Lewis (1969, p. 61), where the claim is made for the particular case of conventions: “to belong to the population in which that convention holds—to be a party to it—is to know, in some sense, that it holds.” A similar, more general claim, can be found in Binmore (1994, p. 140): “[a] community of rational individuals is held together by the pool of common knowledge that I shall call its *culture*.” But for an argument (purely logical) against this view, cf. Gilbert (1989).

<sup>25</sup> Cf. Clark and Marshall (1981, p. 37).

<sup>26</sup> I have argued elsewhere (cf. Sillari 2005) that, with opportune adjustments, the standard mathematical models for knowledge representation *do* accommodate the richer epistemic notions that Lewis uses in his essay.

<sup>27</sup> Cf. Cubitt and Sugden (2003, p. 198).

<sup>28</sup> I am of course referring here to (Aumann 1976).

will be *true*”), or, even if the agent has reason to believe something that is in fact true, her reasons to believe may fail to become a full-fledged, justified (or even *actual*) belief (“there is no assurance that it will be *knowledge*”).

This feature of Lewis’ discussion of common knowledge consists actually in one instance of the well known and rather pervasive problem of logical omniscience<sup>29</sup> in epistemic logic. Briefly stated, the problem of logical omniscience lies in the fact that the characteristics of formal systems devised to mathematically represent knowledge in a group of agents are such that the agents are *perfect reasoners*. For example, they know all the logical consequences of what they know; they know all tautological truths, etc. This is obviously not the case and hence formal epistemic models are inadequate to be a realistic representation of knowledge attributions. In particular, clauses (i) and (ii) above give rise to an infinite series of knowledge attributions, which, of course, cannot realistically be predicated of any material agent.

Lewis states the problem<sup>30</sup> in the first paragraph of the section on common knowledge<sup>31</sup>: “[a]nd how is the process [of generating common knowledge] cut off—as it surely is—so that it produces only expectations of the first few orders?” Lewis’ answer consists in distinguishing the agents’ *reason to believe* and the agents’ *actual belief*, and in fact his definition of “common knowledge” is better characterized as a definition of “common reason to believe.” The infinite hierarchy of knowledge states implicit in the definition of common knowledge is harmless, since it does not consist of any agents’ actual knowledge, but merely of their *potential* beliefs. The distinction is akin to the distinction between implicit and explicit belief made in the literature about *awareness structures*.<sup>32</sup> In our case, we think of reasons to believe as

the operator appropriate to represent the epistemic attitudes of an *ideal* agent<sup>33</sup>, unbounded in computational power and storage capability. Such an agent would believe any logical truth, any logical consequence of what she already knows, etc. On the other hand, a *realistic* agent actually believes only a portion of the infinite amount of formulas that her ideal counterpart would believe. In other terms, whereas the realistic agent has reason to believe an infinite collection of formulas, in actuality she believes only a subset of the collection. According to the intuition that, in order to explicitly know something one must be aware of it, the agent is said to actually believe *p* iff she has reason to believe *p* and she is aware of it.

## 6 Convention and Quantification

In Sect. 4, we have reviewed Cubitt and Sugden’s argument that standard formal epistemic models are inadequate to represent the epistemic notions used by David Lewis in *Convention*, and I argued here and elsewhere (cf. Sillari 2005) that the conclusion drawn in Cubitt and Sugden (2003)—to drop set-theoretical models altogether—is too radical. As we have just seen, using richer epistemic models (technically, non-standard models with awareness) does take into account some of Lewis’ epistemic notions, i.e. the distinction between reason to believe and actual belief. Incorporating the notion of indication involves the addition of both probabilistic reasoning and non-standard structures (cf. for instance Halpern and Pucella 2007; Cozic 2005).

However, there is in Lewis (1969) a further element that traditional epistemic models do not take into account. In fact, a discussion of this portion of Lewis’ epistemic analysis of convention has to the best of my knowledge received little attention in the literature.<sup>34</sup> When analyzing what *kind* of knowledge an agent may have of a convention to which she is a party, Lewis devotes several pages to distinguishing her knowledge of specific instances of the convention, as opposed to her general knowledge of it. Thus, Lewis’ analysis is conducted, at least in part, at the predicate rather than at the propositional level. Analyzing the epistemic status of agents at the predicate level allows,

<sup>29</sup> For a number of technical approaches to deal with logical omniscience in epistemic logic, cf. Fagin et al. (1995). For a philosophical discussion of the problem, cf. Stalnaker (1999, pp. 174–241) and Parikh (1987). Further epistemological reflections on the problem logical omniscience are made in the first part of Sillari (2008a).

<sup>30</sup> Indeed, several authors concerned with presenting *realistic* accounts of common knowledge have addressed the problem, seeking ways to avoid endowing agents with an infinite burden of knowledge attributions. For example, Schiffer, considering possible objections to his definition (cf. Schiffer 1972, p. 36), points out that the definition does not imply that there *is* common knowledge in the group, but only that, by virtue of the conditions in the definition, each agent in the group may *acquire* an indefinite amount of knowledge. Gilbert, in putting forth her own definition of common knowledge (cf. Gilbert 1989, p. 189.), uses the notion of a “smooth reasoning counterpart” of an agent *i*: an ideal agent *i*\* “whose reasoning is untrammelled by limits of time, memory capacity, and perseverance.”

<sup>31</sup> Cf. Lewis (1969, p. 52).

<sup>32</sup> The main reference for awareness structures in the economics literature is Heifetz et al. (2006), while for logic and awareness, allow me to refer the interested reader to Sillari (2008a, b) and the references therein.

<sup>33</sup> For a deep analysis of agency and issues of awareness, cf. Newell (1982), discussed in Sillari (2008a).

<sup>34</sup> Among the authors who have taken into account the distinction, see Burge (1975, p. 251) in which it argued that the “non-triviality” characteristic of a conventional equilibrium need not be common knowledge in either *sensu diviso* or *composito*; or see Loar (1976, pp. 150–151), in which the distinction is applied in philosophy of language to make sense of the fact that an English speaker can have knowledge of so vast and complex a thing as the English language; or see the more recent (Pettit 1998), in which the distinction is called upon to illustrate the original notion of “practical belief” defended in the article.



as we shall see in this section, for important observations that cannot be precisely made within a propositional setup.

What kind of knowledge do parties in a convention share about the convention itself? First of all, notices Lewis, “to belong to the population in which that convention holds—to be a party to it—is to know, in some sense, that it holds.” (Lewis 1969, p. 61.) Thus, *some kind* of knowledge of the convention, maintains Lewis, is a condition to be part of the convention itself (cf. note 22 at p. 12 above). Such knowledge could be “merely potential knowledge” (*ibidem*, p. 63). That is, the agent has, in the sense specified in the previous section, reason to believe that the convention holds, but fails to do the necessary reasoning to achieve conscious, explicit knowledge that the convention holds. Or it could be “irremediably nonverbal knowledge” (*ibidem*), as in the case of Hume’s rowers, who do manage to solve a recurrent coordination problem by synchronizing their movements while they cannot accurately and precisely verbalize a description of their actions.

The kind of knowledge of conventions that interests us, however, is “confined to particular instances” (Lewis 1969, p. 64) rather than to one instance of knowledge “with general content” (*ibidem*). The distinction that Lewis is calling upon is the traditional one between *de re* and *de dicto* knowledge.<sup>35</sup> Of course, both notions of knowledge can be fully general: for the former kind, the agent holds a number of item of knowledge, one for each specific instance; for the latter, the agent holds a unique, general item of knowledge encompassing all particulars. In formal terms, *de re* knowledge is represented by the formula  $K_i \forall x P(x)$ , whereas *de dicto* knowledge is represented by the formula  $\forall x K_i P(x)$ . As Lewis points out, generality *de re* is not straightforward, since each specific instance of the variable  $x$  has to be substituted by some description of an individual, and, of course, substitution in the opaque context of the epistemic operator is problematic.

Besides the traditional problems related to substitutivity in opaque contexts, it is interesting to interpret the distinction cognitively.<sup>36</sup> If the agent’s knowledge of a convention is

knowledge of specific instances only, the agent does not recognize the specific instance at hand as a recurrent one, nor does she understand her action as conformity to the regularity. Notions such as “instance of a recurrent coordination problem  $S$ ” and “behavior conform to a regularity  $R$ ” are not accessible to agents whose knowledge<sup>37</sup> of the existing convention is merely *de re*. To argue that this rather limited kind of knowledge is nevertheless sufficient for the definition of convention (since it allows the agent to “apply his general attitudes to the case at hand,” cf. Lewis 1969, p. 66) Lewis imagines the case of a population of agents that, for some reason, are entirely unable to achieve generality *de dicto* with respect to propositional attitudes and yet can fulfill the clauses of the definition of convention. What seems to be particularly interesting in this analysis is the observation that we who *are* able to generalize beyond specific instances, find ourselves, at times, confined to *de re* knowledge of conventions. The shift from *de re* to *de dicto* generality would then represent the endpoint of some learning process.

Consider the case of Ann and Bob meeting by chance at the cafeteria after the Wednesday seminar, and meeting there the following week, and the following, etc. At the beginning of the semester, they might be unsure as to whether they are meeting every week by chance or purposely. At some point in the semester, it might dawn on them that on *that* particular Wednesday, (i) after the seminar they are going to meet in the cafeteria; (ii) they expect each other to meet there; (iii) they prefer to meet (in the cafeteria) rather than not; (iv) they know that (i)–(iii) are the case. Similarly the following week, and the next one. Thus, a convention in the sense of Lewis is in place. However, Ann and Bob’s knowledge is confined only to specific instances. Later still in the semester, they might come to realize that *in general* after the Wednesday seminar, (i)–(iv) above are the case and full-fledged *de dicto* generality will ensue.

Moreover, Lewis observes (cf. Lewis 1969, p. 66) that knowledge of convention involves not one, but two operations of universal quantification. As we have just seen, one is over instances of the coordination problem, while another is about members of the population. Thus, if Ann is a party to a convention,  $R$  is the conventional regularity and  $S$  is the coordination problem, she expects

<sup>35</sup> In *Convention*, however, Lewis prefers to refer to Abelard’s terminology and talks about knowledge *in sensu diviso* and *in sensu composito*, respectively.

<sup>36</sup> The distinction is akin to different understandings of categorization in cognitive psychology. On the one hand, a category can be seen as defined by a *prototypical* member to which actual category members resemble under certain important criteria. On the other, membership to a category may be decided through the comparison with sets of *exemplars* belonging to the category. As Cristina Bicchieri notices discussing this issue (cf. Bicchieri 2005, p. 84) the use of prototypes, as opposed to that of exemplars, may depend on how well the subject knows the category about which she is making judgments. Thus the distinction between categorization through prototypes and through exemplars reflects some kind of learning process from the former to the latter, like the distinction between *de re* and *de dicto* knowledge of conventions does, as we shall see in this section.

<sup>37</sup> Notice that Lewis’ definition does not simply require that the agents have knowledge of the convention: they also know that they have such knowledge, they know that they know that they have such knowledge, etc. In other words, there has to be a basis  $B$  in the population for common knowledge (or, more precisely, common reason to believe) that the convention holds. But we are considering *de re* knowledge. Thus, the basis  $B$  indicates to the agents that the convention is in place *in each* particular instance of the coordination problem on which the convention is based.

*all* parties to the convention to conform to *R*, in *all* instances *S*. Consider the typical example of a “casual Friday” in the office. Ann—maybe because she is a new hire from a different country—lacks the concept of a “casual Friday”. However, she can see that, in some instances, people in the office are dressing more casually and, in such instances, say, she goes to the locker room during a break to change into more comfortable clothes. She is too shy to inquire with her co-workers about this matter, and hence remains uncertain as to whether the casual code occurs with some regularity in the office. Ann nevertheless suspects that a regularity must be in place (although she does not know which one it might be) and makes various conjectures: is the casual dress code on when the temperature outside is above a certain level? is it on when no especially important clients are expected? is it on when the boss is out of town? And *furthermore*: does it apply only to a certain group of employees? (say, only those with a certain seniority, or only those working in a certain department, etc.) Ann needs to test all such hypotheses which, possibly and in the best case, will all but one be disconfirmed by her experience. At some point, she may consider the hypothesis of a “casual Friday” for the entire staff, and eventually come to realize that it is the correct generalization to make. Although she was complying with the definition of convention even before she realized the nature of the recurrent coordination problem and the extent of the population involved in it, it is only now, having *de dicto* knowledge, that she can appreciate the convention being in place in full generality.

The distinction between *de re* and *de dicto* knowledge of convention seems to replicate some aspects of the distinction between deliberate and automatic abidance by a norm discussed in section 2. Ann has reason to believe, about each instance of the coordination problem, that her part in it is to dress casually. She may be even be aware of her reason to believe, and explicitly know that she is in a coordination problem and that her part in it is to dress casually. And she does dress casually, abiding by the social norm. But, in a sense, her behavior is not rationally justified in full generality unless she recognizes the coordination problem at hand as an instance of the “casual Friday” social norm that is in fact in place, i.e. unless she has general *de dicto* knowledge about the situation at hand. Another important contribution to the understanding of social conventions that can be obtained by considering the *de re/de dicto* distinction is related to the inductive/analogical reasoning involved in the recognition that various instances of coordination problems are in fact similar to each other in important respects and, therefore, are instances of *one* recurrent coordination problem.

## 7 Conclusion

Which, if any, epistemic assumptions are to be included in the definition of social convention? David Lewis’ game-theoretical approach to the problem of defining social kinds such as conventions and norms inescapably presents this question. Of course, the answer is not simple, and I argued in this paper that the answer also needs not be unique. Depending on the approach one takes to attack the problem of explicating the origin and maintenance of social conventions and norms, one may consider the question of epistemic assumptions as fundamental, or one may dispense with it altogether. In any case, it is important to recognize the role that the assumption plays in Lewis’ own account. Common knowledge works primarily, for Lewis, as an instrument of *theoretical* reasoning that justifies the equilibrium beliefs leading the player to conventional coordinative behavior. The Lewisian approach considers repeated games in a given population, yet it is importantly different from today’s evolutionary game theory.<sup>38</sup> The main difference is the presence in Lewis’ semi-formal apparatus of some *cognitive* characteristics of the agents. In Sillari (2005) and in Cubitt and Sugden (2003), two characteristics were singled out and analyzed, namely the distinction between reason to believe and actual belief, and the relation of indication. As we have seen in section 6, the epistemic notions that Lewis uses in *Convention* present a third characteristic distinguishing them from the now standard epistemic formalizations—the distinction between *de re* and *de dicto* knowledge of convention. This distinction may prove useful for investigating issues related to learning and analogical reasoning in the context of social conventions and norms. In turn, a formalization of the rich epistemic notions used by Lewis may prove useful for conducting such an investigation, as it is stressed in Cubitt and Sugden (2003). In Sillari (2005), I argue that awareness structures can formally capture the distinction between reason to believe and actual belief, while in Sillari (2008a, b) I introduce and study a quantified logic of awareness in which the *de re/de dicto* distinction can be taken into account. One of the great challenges stemming from David Lewis’ *Convention* is to recognize and take into account the importance of inductive and analogical reasoning in recurrent coordination problems. The evaluation of the epistemic assumptions made by Lewis, as well as the formalization of the rich epistemic notions involved in them, represent—I hope—important steps towards answering such a challenge.

<sup>38</sup> However, the introduction of cognitive elements in evolutionary analysis—as it is done, for example, in the last chapter of Bicchieri (2005)—is an important point of contact between the two approaches.

**Acknowledgments** The author wishes to thank Cristina Bicchieri, Ken Binmore, Margaret Gilbert, Herbert Gintis, Luca Tummolini, Peter Vanderschraaf and two anonymous referees for valuable comments, suggestions and criticisms.

## References

- Alberucci L, Jaeger G (2004) About cut elimination for common knowledge logics. *Ann Pure Appl Logic* 133(1–3):73–99
- Aumann RJ (1976) Agreeing to disagree. *Ann Stat* 4:1236–1239
- Barwise J (1988) Three views of common knowledge. In: 2nd conference in theoretical aspects of reasoning about knowledge, 1988. Morgan Kaufmann, San Francisco, CA, pp 365–379
- Bicchieri C (1993) *Rationality and coordination*. Cambridge University Press, Cambridge
- Bicchieri C (2007) *The grammar of society*. Cambridge University Press, Cambridge
- Binmore K (1994) *Game theory and the social contract. Playing fair, vol 1*. MIT Press, Cambridge, MA
- Burge T (1975) On knowledge and convention. *Philos Rev* 84(2): 249–255
- Chwe MS-Y (2001) *Rational ritual*. Princeton University Press, Princeton
- Clark HH (1996) *Using language*. Cambridge University Press, Cambridge, MA
- Clark HH, Marshall CR (1981) Definite reference and mutual knowledge. In: Joshi AK, Webber BL, Sag IA (eds), *Elements of discourse understanding*. Cambridge University Press, Cambridge, MA, pp 10–63
- Cozic M (2007) Impossible states at work. Logical omniscience and rational choice. In: Topol R, Walliser B (eds) *Cognitive economics: new trends*, vol. 280, Elsevier, Amsterdam, pp 47–68
- Crawford V, Sobel J (1982) Strategic information transmission. *Econometrica* 50(6):1431–1451
- Cubitt RP, Sugden R (2003) Common knowledge, salience and convention: a reconstruction of David Lewis' game theory. *Econ Philos* 19:175–210
- Fagin R, Halpern J, Moses Y, Vardi M (1995) *Reasoning about knowledge*. The MIT Press, Cambridge, MA
- Geanakoplos J (1992) Common knowledge. *J Econ Perspect* 6:53–82
- Gilbert M (1989) *On social facts*. Princeton University Press, Princeton
- Grice P (1989) *Studies in the ways of words*. Harvard University Press, Cambridge, MA
- Harman G (1977) Review of *linguistic behavior*. *Language* 53(2): 417–424
- Halpern JY, Pucella R (2007) Dealing with logical omniscience. In: *Proceedings of eleventh conference on theoretical aspects of rationality and knowledge*. ACM, New York, pp 169–176
- Heifetz A (1999) Iterative and fixed point common belief. *J Philos Logic* 28:61–79
- Heifetz A, Meier M, Schipper BC (2006) Interactive unawareness. *J Econ Theory* 130:78–94
- Kaneko M, Nagashima T (1996) Game logic and its applications I. *Stud Log* 57:325–354
- Lewis D (1969) *Convention: a philosophical study*. Harvard University Press, Cambridge, MA
- Lewis D (1978) Truth in fiction. *Am Philos Q* 15:37–46
- Loar B (1976) Two theories of meaning. In: Gareth E, John M (eds), *Truth and meaning: essays in semantics*. Oxford University Press, Oxford, pp 138–161
- Meyer JJ-Ch, van der Hoek W (1995) *Epistemic logic for AI and computer science*. Cambridge University Press, Cambridge
- Milgrom P (1981) An axiomatic characterization of common knowledge. *Econometrica* 49(1):219–222
- Newell A (1982) The knowledge level. *Artif Intell* 18(1):87–127
- Parikh R (1987) Knowledge and the problem of logical omniscience. In: *Proceedings of ISMIS- 87 (International Symposium on Methodology for Intelligent Systems)*, pp 432–439
- Pettit P (1998) Practical belief and philosophical theory. *Australas J Philos* 76:15–33
- Schelling T (1960) *The Strategy of Conflict*. Harvard University Press, Cambridge, MA
- Schiffer S (1972) *Meaning*. Oxford Clarendon press, Oxford
- Sillari G (2005) A logical framework for convention. *Synthese* 147(2): 379–400
- Sillari G (2008a) Models of awareness. In: Giacomo B, Wiebe van der H, Michael W (eds), *Logic and the foundations of game and decision theory*, vol. 2 of *Texts in logic and games*. Amsterdam University Press, The Netherlands
- Sillari G (2008b) Quantified logic of awareness and impossible possible worlds. *Rev Symb Log* (forthcoming)
- Skyrms B (1996) *The evolution of the social contract*. Cambridge University Press, Cambridge
- Skyrms B (2003) *The stug hunt and the evolution of the social contract*. Cambridge University Press, Cambridge
- Stalnaker R (1999) *Context and content*. MIT Press, Cambridge, MA
- Sturm H, Wolter F, Zakharyashev M (2002) Common knowledge and quantification. *Econ Theory* 19:157–186
- Sugden R (1986) *The economics of right, co-operation and welfare* [2nd edn., 2004]. Palgrave Macmillan, New York
- Ullman-Margalit E (1977) *The emergence of norms*. Clarendon Press, Oxford
- Vanderschraaf P (1998) Knowledge, equilibrium and convention. *Erkenntnis* 49:337–369
- Vanderschraaf P, Sillari G (2005) Common knowledge. *Stanford encyclopedia of philosophy*. <http://plato.stanford.edu/entries/common-knowledge/>
- Wolter F (1999) First order common knowlegde logics. *Stud Log* 65(2):249–271