Levels of Knowledge, Games, and Group Action

Rohit Parikh
City University of New York[1]
Rohit Parikh

Abstract: In this paper we investigate the possible states of knowledge among a group of individuals and their relations to decision making.

> *A travelling salesman found himself spending the night at home with his wife when one of his trips was accidentally cancelled. The two of them were sound asleep, when in the middle of the night there was a loud knock at the front door. The wife woke up with a start and cried out, "Oh my God! It's my husband!" Whereupon the husband leapt out of bed, ran across the room and jumped out the window.*
>
> Schank and Abelson, 1977, p. 59.

Wimmer and Perner begin their paper [WM] on Beliefs about beliefs with this story from Schank and Abelson which may seem amusing to some and disturbing to others. But the point of the story seems to be that husband and wife each have their own scenario and neither corresponds to the actuality. The wife is a bit better off as she knows *where* she is though not *whom* she is with. The husband is unaware of the identity of his companion and even of his location.

Wimmer and Perner themselves are concerned primarily with the perception by children of other people's *mindsets*. The following quote from [WM] is a story about Maxi which they told a group of children:

> *Mother returns from her shopping trip. She bought chocolate for a cake. Maxi may help her put away the things. He asks her, "Where should I put the chocolate?" "In the blue cupboard," says the mother.*

Later, with Maxi gone out to play, the mother transfers the chocolate from the blue cupboard to the green cupboard. Maxi then comes back from the playground, hungry, and he wants to get some chocolate.

In Wimmer and Perner's experiment, little children who were told the Maxi story were then asked the *belief* question, "Where will Maxi look for the chocolate?"

[1]Department of Computer Science, Brooklyn College, and doctoral programs in Computer Science, Mathematics and Philosophy, CUNY Graduate Center, 365 Fifth Anenue, New York, NY 10016. Email: rparikh@gc.cuny.edu Research supported by grants from the NSF and the CUNY-FRAP program. Some of the results in this paper were presented at *LOFT'02* and at *WOLLIC'02*.

Children at the age of three or less invariably got the answer wrong and assumed that Maxi would look for the chocolate in the *green* cupboard where *they* knew it was. Even children aged four or five had only a one-third chance of correctly answering this question or an analgous question involving Maxi and his brother (who also wants the chocolate and whom Maxi wants to deceive). Children aged six or more were by contrast quite successful in realizing that Maxi would think the chocolate would be in the *blue* cupboard – where he had put it and that if he wanted to deceive his brother, he would lead him towards the green cupboard.

Thus it seems that representation of other people's mindset comes fairly late in childhood, well after they have learned to deal with notions of belief and belief based action for themselves and for others who share their own view of reality. In [St] Chris Steinsvold investigates modal logics which are intended to represent the states of mind of young children. See also [SP].

Older children are not much better. In an experiment in my daughter's seventh grade class, I found that they were unable to deal with the muddy children puzzle beyond the first one or two levels.

In this by now well-known puzzle, a number of children are playing in the mud and some of them get their foreheads dirty. At this the father comes on the scene and announces, "at least one of you has got her forehead dirty."

Scenario 1: Suppose there is only one child, say Amy, who is dirty. Then she will realize that her own forehead must be dirty since she can see that the others are clean.

Scenario 2: Suppose now that there are two dirty children, Sarah and Amy, who are asked in turn, "Do you know if your forehead is dirty?" Now when Sarah is asked, she can see Amy's dirty forehead and she replies, "I don't know." However, when Amy is asked, she is able to reason, "If *my* forehead were clean, Sarah would have known that hers must be dirty since all the others are clean. But Sarah did not know. So my forehead must be dirty."

This reasoning on Amy's part requires a representation by Amy of Sarah's state of mind, and clearly Amy must be at least six for this to work. However, Sarah herself must have some reasoning ability and Amy must know that she has such abilities. It is not enough for Amy to know Sarah's view of reality, she must also represent Sarah's logical abilities in her own mind.

In particular, suppose that there are three dirty children – Jennifer, Sarah, and Amy – who are asked in turn whether they know if they are dirty, and with Amy being asked last. If Sarah is only three, Amy would not be justified in concluding from Sarah's "I don't know" that in that case, Amy herself must be dirty. Amy would need to know that *if* Amy were clean, Sarah would have carried out a representation in her own mind of Jennifer's state of mind and concluded from Jennifer's "I don't know" that Sarah must herself be dirty. But if Sarah is only three, Amy cannot rely on such reasoning on Sarah's part.

As the number of dirty children increases, there is a need for higher and higher levels of "I know that he knows that she knows that...." Common knowledge is at the end of this road and has been offered as the explanation of co-ordinated behaviour ([Lew, HM, CM, Chw]). For instance Halpern and Moses in [HM] show that the co-ordinated attack problem requires common knowledge between the two generals, and that given the means of communication they have, such common knowledge is impossible to attain. Clark and Marshall [CM2] indicate similar difficulties with the referent of "the movie playing at the Roxy today."

While it is true that co-ordinated actions and, supposedly, common knowledge do happen, it may also be relevant to consider other levels of knowledge, short of the infinite, common-knowledge, level.[2] Such levels also arise in certain pragmatic situations, e.g. with e-mail or snailmail or messages left on telephones as voice mail. Thus the purpose of this paper is to study levels other than common knowledge and how they affect the actions of groups.

In typical co-operative situations, even if a certain level of knowledge is needed, a higher level would also do. If Bob wants Ann to pick up the children at 4 PM, it is enough for him to know that she knows. Thus if he sends her e-mail at 2 PM and knows that she always reads hers at 3 PM, he can be satisfied. In such a situation Bob knows that Ann will know about the children in time, or symbolically $K_b(K_a(C))$ and he may feel this is enough. However, if he telephones her at 3 PM instead, this will create common knowledge of C, much *more* than is needed. But no harm done, since in this context, Ann and Bob have the same goals. Halpern and Zuck also state a knowledge level requirement for the sequence transmission problem, which suffices as a minimum, but since the parties are co-ordinating, a higher level does no harm.

But in other contexts one may wish for just a particular level of knowledge, no lower, and no *higher*. Suppose for instance that Bob wants Ann to know about a seminar talk he is giving, just in case she wants to come, but he does not want her to feel pressured to come – she should come only out of interest and not from politeness. In that case he will want to arrange that $K_b(K_a(S))$ (he himself knows that she knows about the seminar), but not $K_a(K_b(K_a(S)))$ (Ann knows that Bob knows that Ann knows about the seminar), for in the latter case she *would* feel pressured. Instead of telling her about his talk, which would create common knowledge, he may arrange for some other method, perhaps for a student to tell her, but without saying that it is a message from Bob.

[2]The following, possibly apocryphal story about the mathematician Norbert Wiener, well known for his absent mindedness, illustrates something even more subtle. At one time the Wieners were moving and in the morning as he was going to work, Mrs. Wiener said to him, "Now don't come home to this address in the evening." And she gave him a piece of paper with the new address. However, in the evening Wiener found himself standing in front of the old address and not knowing what to do – he had already lost the slip of paper with the new address. He went to a little girl standing by and said, "Little girl, do you know where the Wieners have moved to?" The little girl replied, "Daddy, Mom knew what would happen so she sent me to fetch you." The moral of the story, for *us*, is that common knowledge works only if the memory of all parties involved is reliable.

Suppose a pedestrian is crossing the street and sees a car approaching him. It happens in many cities, e.g., Boston, Naples, etc., that the pedestrian will pretend not to notice the car, thereby preventing $K_d K_p(C)$ with C representing the car, d being the driver and p the pedestrian. If the driver knew that the pedestrian knew, he might drive aggressively and try to bully the pedestrian into running or withdrawing. But if he does not know that the pedestrian knows, he will be more cautious.

While the social questions are fascinating and are addressed elsewhere (Cf. [Pa3]), in this paper we shall concentrate on the technical aspects of knowledge, where it is assumed that everyone involved is logically perfect. One can still ask, what are the various levels of knowledge which can arise under various circumstances of communication and how will they affect how we may act?

1 Some examples

Our first example comes from the *Mahabharata*, one of the Indian epics and at 100,000 verses, believed to be the longest single work in the world. It describes the political struggle between two sets of cousins, the *Pandavas*, and the *Kauravas.* In a crucial battle between the two sets of cousins, *Krishna* is an adviser to the *Pandavas*, but out of a sense of fair play he gives his army to the *Kauravas.* At a crucial juncture *Drona*, a powerful warrior on the *Kaurava* side and also the teacher of both the *Pandavas* and *Kauravas*, turns out to be invincible in battle and the *Pandavas* are hard pressed.

However, the wily *Krishna* thinks up a strategem. *Drona's* only son is called *Ashvatthama* and so is an elephant owned by the *Pandavas. Krishna* suggests that the *Pandavas* kill the elephant *Ashvatthama* and then announce that the man *Ashvatthama* has been killed. After a great deal of hesitation and soul searching, self-interest prevails, the *Pandavas* do kill the elephant, and announce to *Drona* that *Ashvatthama* is dead. Drona is a little suspicious, but knows that one of the *Pandava* brothers, *Yudhisthira* never lies. He asks *Yudhisthira*, who confirms that *Ashvatthama* is dead, muttering in an aside, "either man or elephant." Not knowing about the elephant, *Drona* assumes it is his son who is dead, lays down his weapons, and is killed by a warrior on the *Pandava* side.[3]

We now offer a game-knowledge theoretic analysis of this event.

Let *Drona's* two options be f for 'fight' and n for not fight. Before the announcement his preferences were $f > n$ and given his prowess, any warrior who faced him faced death.

But after the announcement that *Ashwatthama* was dead, his preferences change to $n > f$, and he can be attacked with impunity.

Ahswatthama of course was not dead and took terrible revenge on the *Pandavas* includ-

[3]Apparently this half lie was the only untruth ever uttered by *Yudhisthira* and after his death he had to spend a short time in hell – his siblings had longer stays.

ing the killing of some unborn *Pandava* children. For this he was punished by being condemned to live forever, and to wander the earth as a pariah.

Our second example of the motorist and the pedestrian is a bit more complex as the deception consists of inducing a second-order false belief. Let S be the situation where a pedestrian is crossing the street and a car is coming. Let S' be the same situation without the car. In S the pedstrian has two options, g, i.e., to go, and n, i.e., to not go. The motorist also has two similar options, G and N. Here are the payoffs for the two in state S.

Motorist choices

		G	N
Pedestrian choices	g	(-100,-10)	(1,0)
	n	(0,1)	(0,0)

Figure I

Note that there are two Nash equilibria: at (g, N) and at (n, G). However, the penalty for the pedestrian (injury or loss of life) to depart from (n, G) is much greater than the penalty for the motorist (fine or loss of license) to depart from (g, N). Thus the equilibrium (g, N) is less stable than (n, G), and this fact creates the possibility for the motorist to 'bully' the pedestrian.

However, if the pedestrian is unaware of the existence of the car, then the picture is much simpler and his payoffs are 1 for g and 0 for n. g dominates n, and once this choice is made by the pedestrian, it is dominant for the motorist to choose N. This is why the pedestrian tries to achieve the state of knowledge represented by the formulas $K_p(C), \neg K_m(K_p(C))$ indicating that the pedestrian knows the car is there but the motorist does not know that the pedestrian knows. The pedestrian chooses the action g, and knowing that the pedestrian will do this, the motorist must choose N. However, if the motorist has a horn, he can change the knowledge situation. The existence of the car becomes common knowledge and thus the possibility for the motorist to bully the pedestrian arises again.

We now reconsider the problem of the two generals which Halpern and Moses have considered. In this problem there are two generals A, B who are stationed on opposing hilltops and who wish to attack an enemy E in the valley below. General A sends a messenger to general B suggesting that the two should attack together at dawn. However, there is a possibility that the message might not reach B. Perhaps the messenger would be captured or killed by E, and general A attacking by himself would be badly defeated. So general A asks for an acknowledgement from B that his message *let us attack at dawn* has been received.

However, general B has the same problem. He does not wish to attack alone and so he agrees to the attack, but asks for an acknowledgement in turn. Clearly this process of saying *attack at dawn, please acknowledge* can never end. What is needed, Halpern and Moses argue, is common knowledge of the intended attack and no finite number of messages back and forth will achieve it.

However, there is a small twist in this problem to which we now turn.

Let a stand for the event that general A attacks alone.
Let b stand for the event that general B attacks alone.
Let t stand for the event that they attack together.
And finally let n stand for the event that neither attacks.

The analysis which Halpern and Moses provide tacitly assumes that the priorities are: $t > n > b > a$ for general A, and $t > n > a > b$ for general B. "If only one general is going to attack, let it be the other guy," is tacitly assumed. Suppose, however, that $t > b > n > a$ for *both* generals. Perhaps general B's army is large enough to make the difference in a battle, but is also small enough so that its attacking alone and being defeated is not a disaster – in any case it is better than continued inaction. In such a case, general A can issue the message "Let us plan to attack at dawn, please acknowledge, but I will not acknowledge *your* acknowledgement." Thus, A plans to attack iff he receives an acknowledgement, and B has been ordered to attack without worrying about an acknowledgement from A. In such a case, A is guaranteed not to attack alone, and there may well be a good chance of t, provided only that the probability of a message getting through is high enough.

In other words, even though (as Halpern and Moses point out) common knowledge of the attack cannot be achieved, there may well be a strategy for general A which is better than doing nothing.

As our final example we consider the ballot box whose function is to create certain specific states of knowledge. Suppose that five people 1, 2, 3, 4, 5 are debating whether to have lettuce or cucumbers for salad. They cast their votes into a ballot box and the final count reveals that lettuce has won by 3 votes to 2.

Let $C_i : i \leq 5$ mean that i voted for cucumber, and similarly for lettuce. Then the propositional formula which expresses that exactly two voted for lettuce and three for cucumbers is common knowledge. This is a formula with 10 disjunctions, a typical one being $(L_1 \wedge L_2 \wedge C_3 \wedge L_4 \wedge C_5)$. Also what is common knowledge is $(\forall i, \forall j)(i \neq j \rightarrow \neg K_i(C_j))$

as well as $(\forall i, \forall j)(i \neq j \rightarrow \neg K_i(L_j))$. I.e. it is common knowledge that no one knows anyone else's vote. I am sure the reader can see the game theoretic reasons for these two knowledge facts. Everyone must know the results of the election, and the lettuce party must not be in a position to take revenge on the two cucumbers.

We hope we have made a case that states of knowledge and belief which fall short of common knowledge and common belief do arise, are relevant, and can often be better than full common knowledge.

2 Levels of Knowledge

Given a group $N = \{1, ..., n\}$ of agents (whether people or processes), what are the properties of their state of knowledge relative to some fact A? Assuming that A is true, there are still many options. At one extreme, perhaps N have no idea about A. At the other extreme is the possibility that A is common knowledge among N. What are the intermediate possibilities and how will a particular level of knowledge of A affect how the group N will act regarding some situation? To investigate this question formally we set up a formal language L and the notion of the *level* of a formula A as a set of formulas in L.

Definition 1: Assume given m propositional variables $P_1, ..., P_m$. Let $L_0 = \{P_1, ..., P_m\}$ and let L_g be all boolean combinations of the P_i. If the P_i are *basic* ground facts, then L_g represents *all* ground (knowledge-free) facts. Given a group $N = \{1, ..., n\}$ of agents, we define the full knowledge language L as follows:
(i) $L_0 \subseteq L$
(ii) If $A, B \in L$ then so are $\neg A$, $A \vee B$.
(iii) If $A \in L$ then for all $i \leq n$, $K_i(A) \in L$.

To consider common knowledge as well we extend L to L_c by adding the conditions
(iv) $L \subseteq L_c$
(v) If $A \in L_c$ and $U \subseteq N$ then $C_U(A) \in L_c$.
For convenience we shall identify K_i with $C_{\{i\}}$.

Definition 2: A *Kripke structure* $\mathcal{M}$ for L consists of a nonempty set W of states, a map π from $W \times L_0$ into $\{1, 0\}$ with 1 standing for *true* and 0 for *false*, and finally an equivalence relation R_i over W for each $i \leq m$.

Definition 3: Given a Kripke structure $\mathcal{M}$ for L, a state $s \in W$ and a formula $A \in L_c$ we define $\mathcal{M}, s \models A$ as follows by induction on the complexity of A. First, for each $U \subseteq N$, we define the relation R_U to be the transitive closure of $\bigcup R_i : i \in U$. Then we have:
(i) If A is atomic then $\mathcal{M}, s \models A$ iff $\pi(s, A) = 1$
(ii) If $A = \neg B$ then $\mathcal{M}, s \models A$ iff $\mathcal{M}, s \not\models B$
(iii) If $A = B \vee C$ then $\mathcal{M}, s \models A$ iff $\mathcal{M}, s \models B$ or $\mathcal{M}, s \models C$
(iv) If $A = C_a(B)$ where a is either some i or else some U, then $\mathcal{M}, s \models A$ iff

$(\forall t)((s,t) \in R_a \rightarrow \mathcal{M}, t \models B)$

Theorem 1: Let Σ_C be the alphabet whose symbols are $\{C_U\}_{U \subseteq N}$
For all x, y in Σ_C^*, and all formulae A, for all $\mathcal{M}, s$, $V \subseteq U \subseteq N$,
$\mathcal{M}, s \models xC_UC_VyA$ iff $\mathcal{M}, s \models xC_VC_UyA$ iff $\mathcal{M}, s \models xC_UyA$.

In other words, common knowledge by the larger group U absorbs common knowledge by the smaller one.

Corollary 1: Let Σ_K be the alphabet whose symbols are $\{K_1, \ldots, K_n\}$. For all $a = K_i$ in Σ_K, and for all x, y, in Σ_K^*, and all formulae A,

$$\models xayA \leftrightarrow xaayA$$

and hence for all $\mathcal{M}, s$, $\mathcal{M}, s \models xayA$ iff $\mathcal{M}, s \models xaayA$, i.e., repeated occurrences of a are without effect and if $xay \in L_K(A,s)$ then $\forall n\ xa^ny \in L_K(A,s)$.

In other words, it is common knowledge that a knowing some B is the same as a knowing that a knows B.

Definition 4: Given a formula A and $\mathcal{M}, s$ the *level* of A at s, $L(A,s)$ is the set of x in Σ_C^* such that $\mathcal{M}, s \models xA$, and x contains no substrings C_UC_V, C_VC_U for any $V \subseteq U \subseteq N$.

Strings x such that x contains no substrings C_UC_V, C_VC_U for any $V \subseteq U \subseteq N$ will be called *simple*, and from now on we shall confine ourselves to simple strings.

If s is clear from the context, or not important, then we shall drop it as a parameter. If we restrict ourselves to the K_i operators, we denote the level of A at s by $L_K(A,s)$.

3 Embeddability

Now we will try to characterize levels of knowledge. First we need to introduce the embeddability ordering on strings which turns out to be important here.

Definition 5: Given two strings $x, y \in \Sigma_K^*$, we say that x is *embeddable* in y ($x \leq y$), if all the symbols of x occur in y, in the same order, but not necessarily consecutively. Formally:
1) $x \leq x$, $\epsilon \leq x$ for all x
2) $x \leq y$ if there exist x', x'', y', y'', such that $x = x'x''$, $y = y'y''$, and $x' \leq y'$, $x'' \leq y''$.
and $\leq$ is the smallest relation satisfying (1) and (2).

Thus the string *aba* is embeddable in itself, in *aaba* and in *abca*, but not in *aabb*.

Properties of the embeddability relation $\leq$

Fact 1: Embeddability is a *well-partial order*, i.e. it is not only well-founded, but every linear order that extends it is a well-order. Equivalently, it is well-founded and every set of mutually incomparable elements is finite.

Note for instance that an infinite set of incomparable elements $\{a_1, ..., a_n, ...\}$ is well-founded – nothing is below anything else. However, it is not a WPO, for we can clearly set $a_1 > a_2 > a_3 ....$ which gives an extension of the original, flat ordering. The flat ordering was well-founded, but the extension is not. Thus the condition that an ordering be a WPO is much stronger than the condition that it be well-founded. Note that any extension of a WPO will also be a WPO. For if $< \subseteq <'$ and $<$ is a WPO, then every linear extension of $<'$ is an extension of $<$ and hence a well-order. Thus $<'$ is a WPO.

Fact 2: Embeddability can be tested in linear time, e.g., by a nondeterministic finite automaton with two input tapes.

Fact 1 was proved first by Graham Higman [Hi]. See [JP] for a discussion. Fact 2 is straightforward.

We also need a stronger relation defined on Σ_C^*, which we call *C–embeddability*.

Definition 6: Given two strings $x, y \in \Sigma_C^*$, we say that x is *C–embeddable* in y ($x \preceq y$), if
1) If $V \subseteq U$ then $C_V \preceq C_U$
2) $x \preceq y$ if there exist x', x'', y', y'', $(y', y'' \neq \epsilon)$, such that $x = x'x''$, $y = y'y''$, and $x' \preceq y'$, $x'' \preceq y''$.
and $\preceq$ is the smallest relation satisfying (1) and (2).

Fact 3: For any $x, y \in \Sigma_K^*$, $x \leq y$ iff $x \preceq y$.

Fact 4: C-embeddability is a well-partial order.

Proof: Fact 3 is easy. It is also easy to check that C-embeddability is a partial order. It is well-founded, because regular embeddability is well-founded and for given $x \in \Sigma_C^*$ there are only finitely many $y \in \Sigma_C^*$ s.t. $|x| = |y|$ and $y \preceq x$.

To see that it is a WPO, consider dropping the condition that if $V \subseteq U$ then $C_V \preceq C_U$. We would then get a WPO, for we would essentially be treating each C_U as an independent symbol, unrelated to any other C_V. We are looking at the embeddability ordering on an alphabet of 2^n elements where n is the number of agents. By the observation earlier, we have a WPO. Adding condition (1) gives us an extension of a WPO which is therefore also a WPO.

There are only finitely many incomparable elements in Σ_C^* with respect to $\leq$, and there are more incomparable elements with respect to $\leq$ than with respect to $\preceq$, so $\preceq$ is a well-partial order. □

If $\leq$ is a partial order on S, we can define a notion of a *downward closed* subset of S:

Definition 7: $R \subseteq S$ is *downward closed* iff $x \in R$ implies $\forall y \preceq x, \quad y \in R$.

4 The Main Results on Levels of Knowledge

The following result about levels of knowledge (with or without common knowledge) follows from theorem 1.

Theorem 2: Let Σ_C be the alphabet whose symbols are $\{C_U\}_{U \subseteq N}$. Then for all strings x, y in Σ_C^*, if $x \preceq y$ then for all $\mathcal{M}, s$, if $\mathcal{M}, s \models yA$ then $\mathcal{M}, s \models xA$.

Corollary 1: Every level of knowledge is a downward closed set with respect to $\preceq$. □

Theorem 3: There are only countably many levels of knowledge and in fact all of them are regular subsets of Σ^* (where Σ is either Σ_K or Σ_C).

Proof: Let L be any downward closed set of some Σ^*. Let $X = \Sigma^* - L$. Then X is upward closed. But now let M be the set of *minimal* elements of X. Since embeddability is a WPO, M is finite and let $M = \{x_1, ..., x_p\}$. Then for all $y \in \Sigma^*$, $y \in X \leftrightarrow (\exists i \leq p)(x_i \preceq y)$ and the condition $(\exists i \leq p)(x_i \preceq y)$ can be tested by a finite automaton. Thus X is regular, and hence so is L. □

Fact 5: Eric Pacuit of the CUNY Graduate center and ourselves have shown that in contrast with *knowledge* there are *uncountably* many possible levels of rational *belief*. This is curious, as truth is the only condition which (formally) separates knowledge from rational belief. These results will appear elsewhere.

Corollary 1: The membership problem for a level of knowledge can be solved in linear time.

Now we consider what finite downward closed sets of strings can look like.

Theorem 4: If L is a non-empty finite subset of Σ_K^*, then L is downward closed iff for some k,

$$L = \bigcup_{i=1}^{k} dc(\{x_i\})$$

where $x_i \in \Sigma_K^*$.

Proof: Consider the set M of *maximal* elements of L. Then because the order is a WPO, the set M must be finite. Moreover, every element of L must lie below some maximal element. Hence if $M = \{x_1, ..., x_k\}$ then we get $L = \bigcup_{i=1}^{k} dc(\{x_i\})$. □

This theorem reiterates the fact that the finite levels are characterized by their maximal elements ($x_1, ..., x_k$ are maximal). The characterization of infinite levels of knowledge is more complex. The details are in [PK].

We have shown that every level of knowledge is a regular set of strings satisfying certain conditions. But do all such sets actually arise as levels of knowledge? We now give a simple argument to show that they do in fact. The following result is proved jointly with

Eric Pacuit.

Theorem 5: Let L be a downward closed set of strings relative to $\preceq$. Then there is a finite Kripke model M and state s such that for all strings x in $\{K_1, ..., K_n\}^*$, $M, s \models xp$ iff $x \in L$, where p is a propositional variable.

Proof: We assume that L is not empty, for otherwise we can simply achieve the desired effect by making p false at s. Since L is not empty and is downward closed, it follows that the empty string belongs to L. Also as we saw earlier, there are finitely many strings $y_1, ..., y_k$ which are mutually incomparable and such that $x \in L$ iff there is no i such that $y_i \preceq x$. We can assume that the strings y_i are without repetitions as repetitions can be removed without any harm.

Let the base set W of the model consist of the set of all s_y, where y is a simple string and $y \preceq y_i$ for some i. We make p true at s_y if $y \neq y_i$ for any i and p false at s_y if $y = y_i$ for some i. Moreover, for each pair of strings x, y where $y = xK_i$ or $y = x$ we let $(x, y), (y, x) \in R_i$, the accessibility relation corresponding to agent i. Even though each R_i is an equivalence relation, the equivalence classes will have at most two elements.

We note the following. If there is a sequence of states $s_0 = s, s_1, ...s_m = s_x$, and for each $j < m, (s_j, s_{j+1}) \in R_{i_j}$, then the string x is embeddable in $K_{i_1}....K_{i_m}$. Because of the symmetry of each R_i, backward movement $s_j \to s_{j+1} \to s_j$ is possible, so m can be larger than the length of x. But that does not vitiate the claim.

We now show that $M, s \models xp$ iff $x \in L$. So suppose not, and let $x = K_{i_1}K_{i_2}...K_{i_m}$. For xp to fail at s there must be a chain of states $s = s_1, ..., s_{m+1}$ such that for each $j \leq m$ $(s_j, s_{j+1}) \in R_{i_j}$ and moreover $M, s_{m+1} \models \neg p$. Then we must have $s_{m+1} = s_y$ for some $y = y_k$. But then the string y_k is embeddable in x and hence $x \notin L$.

On the other hand if $y_i \preceq x$ for some i, then let for example $y_i = K_1K_2K_1$ and $x = K_1K_2K_3K_1$. Since the relations R_i are all reflexive, there is a path $s_1, ..., s_5$ from $s = s_1$ upto $s_5 = s_x$ such that $(s_1, s_2) \in R_1, (s_2, s_3) \in R_2$, etc till $(s_4, s_5) \in R_1$. Hence $M, s \models \neg(K_1K_2K_3K_1p)$, i.e., $M, s \not\models K_1K_2K_3K_1p$. Thus $M, s \models xp$ iff $x \in L$. □

In [PK] we describe the sort of levels of knowledge which can arise in concrete models created by distributed processes. Only finite levels of knowledge can be created by asynchronous communication, and n-person broadcasts cannot create common knowledge of any non-trivial formula among $n + 1$ agents. Other than that, most levels are attainable under suitable scenarios.

5 Applications to Games

In the following section we investigate the connection between game theoretic strategies and levels of knowledge. We will confine our discussions to two-player games.

Following the examples earlier (*Mahabharata*, pedestrian, etc.) we assume that there are two payoff matrices[4] M, N. To simplify matters we assume that the strategies are the same for both matrices, so they are the same size; only the payoffs are different. We also assume that N is the *default* matrix, i.e., that both players assume that in the absence of other information, the payoffs are according to N. However, the actual matrix is M and p is the proposition that M is the actual matrix. The payoff for the two players in M are $r(i, j), c(i, j)$ respectively if row i and column j are played. The payoffs in N are $r'(i, j), c'(i, j)$.

Clearly if neither player is aware of p they will both play according to N. Let us make the further assumption that in the default situation, there is a dominant strategy for the column player C. Without loss of generality we assume this to be column 1. We will number the strategies of the column player by odd numbers and those of the row player by even numbers.

Suppose now that R knows p and C does not and R knows that as well. In that case R can safely assume that C will play 1, and so R will play that strategy i which maximizes $r(i, 1)$. Suppose it is strategy 2. Then the payoffs to the two players will be $r(2, 1), c(2, 1)$, even though player C had expected something of the form $c'(i, 1)$.

Let us take this analysis one step further. Suppose now that we have $K_r(p), K_c(K_r(p))$, $K_c(\neg K_r(K_c(p))$. So R knows that p but does not know that C knows this. R will play strategy 2 as above, and knowing this, C will play the best response to this, say strategy 3. Thus their payoffs will be $r(2, 3), c(2, 3)$ respectively, even though R had expected to get $r(2, 1)$. Will R be unpleasantly surprised? Not necessarily, it depends on the kind of game. If there are only two plays and the game M is one of co-ordination, i.e. for all i, j, k, l, $c(i, j) > c(k, l)$ iff $r(i, j) > r(k, l)$ then what is to one player's benefit is also to the benefit of the other.[5]

Game G: This is really a pair of games, but having warned the reader we will just call it a game. In this game each player chooses a number between 1 and 10. If the two numbers are more than 1 apart, the payoff is 0 in both M, N. If the numbers chosen are at most 1 apart, then in N the payoff to each player is the *maximum* of the two numbers, say a, b. However, if they are at most 1 apart, then the payoff in M will be 10 - min(a, b). So in N it pays to pick the higher numbers, and in M it pays to pick the lower numbers. Thus for instance the numbers (2,6) will yield 0 payoff in both matrices, whereas (2,3) will yield a payoff of 3 in N and of 8 in M.

Now if C does not know p, C – assuming that the matrix is N – will play 10. In N, this would give the highest possible payoff of 10, provided that R plays 9 or 10. Now if R knows p and that C does not know p, R will know that C will play 10 and R herself will play 9. Thus the payoff to each will be $10 - 9 = 1$.

[4]The game between row and column with two matrices M, N could in fact be thought of as a single game with three players.

[5]Co-ordinated games are a special case of *Games of Common Interest* considered by Aumann and Sorin [AS].

Suppose now that C knows that R knows p, but R does not know this. Then C will know that R will play 9 and C will play 8, with the payoff to both being 2. R will be surprised, but pleasantly.

As the level of knowledge of p goes up, so will the payoffs, until a maximum of 10 is reached with one of the players playing 0 and the other playing 1. Even though the payoffs are co-ordinated, higher levels of knowledge bring greater benefits, until level 10 is reached and after that, even common knowledge will be no better.

In any case, it is evident that in finite strategy games with co-ordinated payoffs, the higher levels of knowledge always bring better payoffs. This will be the case in two player games of co-ordination, provided only that one player knows everything which the other does.

The situation is different if there are more than two players or if there are infinitely many strategies for both players. With more than two players a problem can arise if players have incomparable knowledge. Thus if there are three players A, B, C, and A and B have different notions of what C will play, then even though the three have common interests, A and B might make choices which will make the outcome worse for all three. But if there is a hierarchy so that player A knows everything which B knows, and player B knows everything which C knows, *and* the games are co-ordinated so that a benefit to one is a benefit to the others, then we will still have the result that higher, *finite* levels are beneficial.[6]

Game G2: In this game both players play some natural number. The payoff is 0 in both M, N if the difference between the two numbers is more than 1. In N, if either number exceeds 10, the payoff is 0, but if both numbers are ≤ 10 and no more than 1 apart, the payoff is the maximum of the two numbers. Matrix M is similar to matrix N in that higher numbers are better, but there is no punishment for numbers > 10. The payoff is the maximum, period, provided only that the numbers are no more than 1 apart.

Suppose now that p is true, R knows it but C does not. Then C will play 10 and knowing this, R will play 11. The payoff will be 11 for both.

If $K_r(p), K_c(K_r(p)), K_c(\neg K_r(K_c(p)))$, then C will know that R will play 11 and C himself will play 12, thus getting a payoff of 12 for both. As the level of knowledge rises, so will the payoff. But now there is a paradox! If p is common knowledge, the players will have *no* idea how to play! So common knowledge is not necessarily better than a high finite level of knowledge.

Non-coordinated games: Suppose now that we are dealing with games where the payoffs are not positively correlated. Perhaps one or both matrices are zero sum, although as we noticed, it is really the matrix M which counts. Matrix N is only used to establish that 1 is the default strategy for C.

[6]Perhaps this technical result substantiates something which we all know, but do not like, that in situations where there is common interest, hierarchies can be beneficial.

Now suppose we have some level of knowledge $K_r(p), K_c(K_r(p)), ...$ rising to some finite level. Then we may have the situation that the default play for C (in N) is 1, the best response to that (in M, the actual game) is 2, the best response to that is 3 and so on, until we reach the play corresponding to the actual level of knowledge, perhaps $(r(2n, 2n+1), c(2n, 2n+1))$ for some n.

Now if the matrices are finite, then there is bound to be a cycle. Perhaps such a cycle will reach a Nash equilibrium and then stabilize. Perhaps it will just go on and on, with immediately succeeding levels of knowledge giving different payoffs, with the player who knows a bit more having an advantage.

6 Further Work and Open Questions

In this paper we have looked only at levels of knowledge for single formulas. However, levels of knowledge for related formulas may be connected. For example, if $A = B \vee C$, then $L(B, s) \cup L(C, s) \subseteq L(A, s)$. So one could ask, given the Lindenbaum algebra $\mathcal{A}$ of ground formulas and the Boolean algebra $\mathcal{B}$ of subsets of Σ_c^*, which maps from $\mathcal{A}$ to $\mathcal{B}$ can arise as level maps? We know that the maps must preserve order and that the images must be regular, downward closed sets, but what more can we show?

A second direction of inquiry is to ask how actual game playing and knowledge interact. We have shown what sorts of levels can arise and shown that they are relevant to group strategies as well as to individual strategies within groups. But clearly much more needs to be done.

A final line of research is to bring the current work into closer contact with a lot of other work on knowledge revision which begins with Plaza and proceeds through [Ger], [BMP], [Dit]. We also need to relate the work with the work of Stalnaker on models of knowledge where probabilities are taken into account.

Acknowledgements: We thank Steven Brams, Eva Cogan, Joe Halpern, Karen Kletter, Eric Pacuit, Debraj Ray and two referees for very useful comments.

References:

[AS] R. Aumann and S. Sorin, Cooperation and bounded recall, *Games and Economic Behavior* **1** (1989) 5-39.
[Ba] J. Barwise, Three views of common knowledge, in *TARK-2*, Ed. M. Vardi, Morgan Kaufmann 1988, pp. 369-380.
[BMP] A. Baltag, L. Moss and S. Solecki, The logic of public announcements, common knowledge and private suspicions, Technical report #238, Indiana University Cognitive Science Program (earlier version appeared in *TARK 1998*)
[Chw] M. Chwe, *Rational Ritual*, Princeton U. Press, 2001
[CM2] H. H. Clark and C. R. Marshall, Definite reference and mutual knowledge, in *El-*

ements of Discourse Understanding, Ed. Joshi, Webber and Sag, Cambridge U. Press, 1981.
[Dit] H. van Ditmarsch, *Knowledge Games*, Doctoral dissertation, U. of Groeningen (2000).
[FHMV] R. Fagin, J. Halpern, Y.Moses and M. Vardi, *Reasoning abour Knowledge*, MIT press, 1995.
[Ger] J. Gerbrandy, Dynamic epistemic logic, in *Logic, Language and Information II*, CSLI press 1999.
[Hi] G. Higman, Ordering by divisibility in abstract algebras, *Proc. London Math. Soc.* **3** (1952) 326-336
[HM] J. Halpern and Y. Moses, Knowledge and Commmon Knowledge in a Distributed Environment, *Proc. 3rd ACM Symposium on Distributed Computing* 1984 pp. 50-61
[HZ] J. Halpern and L. Zuck, A little knowledge goes a long way, *Proc. 6th PODC*, 1987, pp. 269-280.
[JP] D.H.J. de Jongh and R. Parikh, Well partial orderings and hierarchies, *Proc. Kon. Ned. Akad. Sci* Series A 80 (1977) 195- 207.
[Lew] D. Lewis, *Convention, a Philosophical Study*, Harvard U. Press, 1969.
[MGM] R. Marvin, M. Greenberg and D. Mossler, The Early development of conceptual perspective thinking, *Child Development*, **47** (1976) 511-514.
[MT] Y. Moses and M. Tuttle, Programming simultaneous actions using common knowledge, Research Report MIT/LCS/TR-369 (1987)
[Pa1] R. Parikh, Knowledge and the problem of logical omniscience, *ISMIS-87*, North Holland, pp. 432-439.
[Pa2] R. Parikh, Finite and infinite dialogues, *Proceedings of a Workshop on Logic and Computer Science*, ed. Moschovakis, Sprinter 1991, 481-98.
[Pa3] R. Parikh, Social software, *Synthese*, **132**, Sep 2002, 187-211.
[PK] R. Parikh and P. Krasucki, Levels of knowledge in distributed computing, *Sadhana – Proc. Ind. Acad. Sci.* **17** (1992) pp. 167-191.
[Pl] J. Plaza, Logics of public announcements, *Proceedings 4th International Symposium on Methodologies for Intelligent Systems*, 1989.
[PR] R. Parikh and R. Ramanujam, Distributed computing and the logic of knowledge, *Logics of Programs* 1985, Springer LNCS 193, 256-268.
[SA] R. Schank and R. Abelson, *Scripts, Plans, Goals, and Understanding*, Erlbaum Hillsdale, NJ (1977)
[SP] C. Steinsvold and R. Parikh, A Modal analysis of some phenomena in child psychology, *Bulletin of Symbolic Logic*, Mar 2002, Logic Colloquium '01, page 158.
[Sta] R. Stalnaker, Knowledge, belief and counterfactual reasoning in games, in *The Logic of Strategy*, Ed. Bicchieri et al, Oxford University Press, 1999.
[St] C. Steinsvold, Trust and other modal phenomena, research report, CUNY Graduate Center, February 2002.
[WP] H. Wimmer and J. Perner, Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception, *Cognition*, **13** (1983) 103-128.