Knowledge games

Submission to the Bulletin of Economic Research

Hans van Ditmarsch*†‡

November 25, 2000

Abstract

The subject of epistemic logic is firmly entrenched in game theory, including the analysis of common knowledge and of public announcements, such as in 'hat problems' [OR94, Bin92], see also [FHMV95]. How to analyse communications to *sub*groups of the public, and the effects of such common knowledge of a subgroup on the information state of a larger group, has only recently come into fruition [Ger99, Bal99, vD00]. Knowledge games are introduced to provide a comfortably concrete vehicle for the study of such interactions. We introduce the concepts of knowledge game, deal of cards, knowledge game state, game action, and action execution. A deal of cards is a function from cards to players. A knowledge game state is represented by a pointed multiagent $S5$ model on the set of card deals where all players hold the same number of cards as in the actual deal. A game action combines a question with an answer, and is represented by a pointed multiagent $S5$ frame on the set of possible answers. The execution of a game action in a knowledge game state corresponds to the computation of a pointed multiagent $S5$ model that is a restriction of the direct product of the corresponding action frame and game model.

1 Introduction

Imagine a country mansion with a couple of partying guests. Suddenly the host is discovered, lying in the basement, and murdered. The guests decide to find out among themselves who committed the murder.

The body is discovered by the butler, under suspicious circumstances that indicate that the location is not the actual murder room. In order to solve the murder it is required to find out who the murderer is, what the murder weapon was, and in which room the murder was committed. The butler is exonerated, the six guests are therefore the suspects. The guests are: Colonel

*University of Groningen, Cognitive Science and Engineering, Grote Kruisstraat 2/1, 9712 TS Groningen, the Netherlands, hans@cs.rug.nl, http://tcw2.ppsw.rug.nl/~hans/

†Figure 1 is reprinted with permission by Hasbro.

‡I thank the anonymous referee for his/her useful comments.

Figure 1: Examples of a guest, a weapon and a room card

Mustard (colour yellow), Professor Plum (colour pink), the Reverend Green, Mrs. Peacock (colour blue), Ms. Scarlett (colour red, i.e. scarlet), and Mrs. White. There are six possible murder weapons: candlestick, rope, leaden pipe, wrench, gun, knife. The house consists of nine different rooms: hall, kitchen, dining room, study, sitting room, patio, ballroom, library, pool room.

The game consists of a board with a picture of the house, with the nine rooms in it and 'paths' leading in a certain number of steps from one room to another. Also there are six suspect cards, six weapon cards, nine room cards. A pair of dice, six pawns for the (six) players, in colours matching the guests' names, and six weapon tokens complete the picture.

There are six players. The three types of cards are shuffled separately. One suspect card, one weapon card and one room card are blindly drawn and put apart. These 'murder cards' represent the actual murderer, the murder weapon and the murder room. All remaining cards are shuffled together. They are then dealt to the players. Every player gets three cards. Some player starts the game, which is determined by throwing dice, or by the general rule that the player with the red pawn starts. That player then makes a move. A move consists of: throwing the dice; trying to reach a room by walking your pawn over the game board (the number of steps on the board may not exceed the outcome of the throw of dice); *if* a room is reached voicing a *suspicion* about it, i.e. about a suspect, a weapon and that particular room; gathering responses to that suspicion from the other players; and optional: making an *accusation* about a suspect, a weapon and a room.

As a consequence of the suspicion, the pawn with the same colour as that of the suspected player is moved to the suspected room. The other players are supposed to respond to the suspicion in clockwise fashion: the first player that holds at least one of the three cards mentioned in the suspicion, must show exactly one of those to the requesting player, and to him only. This ends the move. Whoever is next in turn is again determined clockwise. Just like a suspicion, also an *accusation* is the combination of a suspect, a weapon and a room card. Each player can make an accusation only once in the game. It is not voiced but written down. The accusing player then checks the three murder cards, without showing them to others. If the accusation is false, that player has lost and the game continues. The first player who correctly guesses the murder cards, wins the game. Note that, although pawns are identied with guests, you don't even know whether you have committed the murder 'yourself'.

Given some simplifications, the game of Cluedo is nothing but a game where a finite number of cards (in this case 21) are dealt over a finite number of players (in this case 7: six 'real' players and the table), and where actions consist of either questions and answers about cards, or are announcements about (not) winning. A deal of cards is a *function* from cards to players. We will show that game states, game actions, and the transitions resulting from their execution can, amazingly, *all* be defined by operations on the function space of deals of a given number of cards to a given number of players.

We call Cluedo and similar games *knowledge games.* A knowledge game is informally defined by a deal of cards over players, a set of possible game actions (or moves), an order protocol to determine who is to move next, and a procedure to determine who wins. Cards do not change hands during a play of the game, although they may be shown. Players know their own cards and know how many cards all players have. The state of the game is fully determined by the deal of cards and by the game action sequence, initially empty. Although cards do not change hands, *knowledge* about cards does change during the game, and only that: therefore we have named these games *knowledge* games.

In section 2 we give an example of a very simple knowledge game, played by three players with each one card. In section 3 we discuss deals of cards. In section 4 we define the state of a game, as a pointed multiagent $S5$ model. In section 5 we define game actions. In section 6 we define the execution of a game action in a knowledge game state. In section 7 we conclude with some tentative remarks on the *game* aspect of knowledge games.

In this article we restrict ourselves to a purely structural and operational analysis of game states and game actions. Both game states and game actions have (unrelated) peculiarities so that their detailed logical description can be regarded as a contribution to epistemic logic. In [vD00] we characterize initial knowledge game states in a multiagent epistemic logical language. Also in [vD00], we define a (general purpose) language with dynamic modal operators for 'knowledge actions' that describe game actions. Because of this separate treatment, familiarity with epistemic logic is *not* required for reading this article.

Our main contribution to game theory is that we show that plays of card games can be *entirely* described in an epistemic setting. Secondly, within this setting we give examples where growth of knowledge can *not* be seen as refining partitions (information sets) on a given set of states, i.e. as eliminating possibilities. The last appears to be an implicit suggestion in the literature [Bin92]. We show that the set of states may increase, as they incorporate players' changing beliefs about each other, so that we may no longer speak of refinement.

The analysis of game states and game actions is a prerequisite to computing optimal strategies for knowledge games (see section 7). We certainly regret that we have not yet continued our research in that direction, but hope that within its current restrictions the subject remains of sufficient interest to the economist.

2 Three players and three cards

Even when reduced to a knowledge game, the game of Cluedo is rather complex. We start by giving an example of a simpler knowledge game. The game for three players each holding a card, is the simplest kind of knowledge game that still contains most of the features that we consider interesting.

Example 1 (The hexa game)
Consider the following game. There are three players. They are called 1, 2 and 3. There are three cards. The cards are called red, white and blue, or r, w, b (the colours of the Dutch flag). Every player is holding one card. Players can only see their own cards. A player can ask a question to another player. The question should always be answered. Also, after the question has been answered, the requesting player may announce that he knows what the deal of cards is. The first player to do so, wins the game. Players never lie, are perfect reasoners, know the kind of game they are playing, etc. We call this game the hexa game.

Only some kinds of question are permitted. A player can ask another player for one particular card, or for one of two cards, or for for one of three (one of all) cards. A question for one of three cards is a question for 'his card'. It is a rule of the game, for how many cards one can ask. In other words: you cannot choose between asking for one card or for two cards, depending on how informative you expect an answer to be.

Suppose the actual deal of cards is: 1 holds red, 2 holds white, and 3 holds blue. We call that deal: rwb. Suppose player 2 asks player 1 "do you have the red or the blue card?". Given that the actual deal of cards is rwb, we can imagine player 1 to respond by saying "yes" (i), by privately showing player 2 the red card (ii), by only showing player 2 the red card (iii), or by publicly (face up) showing player 2 the red card (iv). In (ii), by 'privately' we mean that player 3 is not aware of (and does not suspect) the question being answered. In the resulting state player 3 incorrectly 'knows' that 2 does not know that 1 holds red. Therefore, answer (ii) makes no sense in a *game*. In (iii), by 'only' we mean that player 3 is seeing that a card is being shown, and that 1 and 2 know that he is seeing it, etc., but that 3 cannot see that it is the red card. In knowledge games we permit only answer (iii). Answer (iv) is also equivalent to *saying* "yes, namely the red card".

Apart from showing a card, there is one other type of answer to a card request. If, given deal rwb, player 2 asks player 1: "do you have the blue card", player 1 says: "no". Thus we allow two types of answer to a request: showing a requested card to the requesting player, and to him only; or saying that you do not have any of the requested cards. The combination of a request with an answer is a *game action.*

We now play an entire game. The deal of cards is rwb. The questions must be for one of three cards (a card). Player 2 starts. Player 2 asks player 1 for his card. Player 1 shows player 2 the red card. Player 2 says that he knows the deal of cards. Player 2 has won. Player 2 cannot lose this game.

We play again. The questions must be for one of two cards. Player 2 starts. Player 2 asks player 1 for the red or the white card. Player 1 shows player 2 the red card. Player 2 says that he knows the deal of cards. Player 2 has won. Player 2 cannot lose this game.

We play again. The questions must be for one card. Player 2 starts. Player 2 asks player 1 for the red card. Player 1 shows player 2 the red card. Player 2 says that he knows the deal of cards. Player 2 has won. Player 2 could have lost this game by playing differently: Player 2 starts. Player 2 asks player 1 for the white card.[1] Player 1 answers 'no, I don't have it'. Player 2 ends his move. It is now the turn of player 3. Player 3 says that he knows the deal of cards. Player 3 has won.

3 Deal of cards

We continue by introducing relevant concepts for knowledge games.

Definition 1 (Deal of cards)
A *deal* is a function $d : \mathbf{C} \to \mathbf{A}$ from a finite set $\mathbf{C}$ of cards to a finite set $\mathbf{A}$ of players or agents.

We can think of each player a holding the cards $c \in d^{-1}(a)$. Observe that some players may hold zero cards. We generally name the players $\mathbf{A} = \{1, 2, ..., n\}$, and the cards $\mathbf{C}$ with lower case letters. We sometimes distinguish a nonactive player, the cards on the table so to speak. In that case we assume $0 \in \mathbf{A}$ and take player 0 to be the 'table'.

We assume that deals are *total* functions, i.e. that all cards have been dealt. This is without loss of generality: suppose a deal d were partial, so that some cards are not dealt to any player, and, so to speak, remain in the stack of cards. We assume these 'remaining cards' to be 'on the table', i.e. they are held by the imaginary player 0.

Notation Let d be the deal of three cards over three players such that 1 holds red, 2 holds white, and 3 holds blue; thus $d(r) = 1$, $d(w) = 2$, $d(b) = 3$. We introduce a shorthand notations for deals. We leave the players implicit and list only the cards they hold, assuming the numerical order of players, (possibly)

[1] Although players are perfectly logical, we do not require them to be perfectly rational: they know all the deductive consequences of answers to their questions, but they cannot justify preferences among questions.

separated by vertical bars. In this case we get $r|w|b$, or simply rwb, as above. If a player doesn't hold any cards, write ε. If player 3 holds a fourth, yellow (y) card as well, we get $r|w|by$ (or $rwby$). If, instead, player 2 holds no card at all, we get $r|\varepsilon|b$ (or $r\varepsilon b$).

Definition 2 (Size of a deal of cards)
Let $d \in \mathbf{A}^{\mathbf{C}}$. Write $|\mathbf{A}| = n$. The *size of deal* d, notation $\sharp d$, lists for each player the number of cards he holds:

$$\sharp d := |d^{-1}(1)| \mid ... \mid |d^{-1}(n)|$$

Deals where all players hold the same number of cards are said to be of the same size. Unless confusion results, we delete the vertical bars of separation and write $|d^{-1}(1)|...|d^{-1}(n)|$. Thus $\sharp(r|w|b) = 1|1|1$ (or 111), $\sharp(r|w|by) = 1|1|2$ (or 112), etc.

Definition 3 (Set of deals of the same size)
Given a deal $d \in \mathbf{A}^{\mathbf{C}}$, $D_{\sharp d}$ is the set of deals of the same size as d (the set of deals where all players hold the same number of cards as in d):

$$D_{\sharp d} := \{e \in \mathbf{A}^{\mathbf{C}} \mid \sharp d = \sharp e\}$$

We assume that players see their own cards, and see how many cards every other player holds. This induces an equivalence relation on $\mathbf{A}^{\mathbf{C}}$.

Definition 4 (Accessibility between deals of cards)
Let $a \in \mathbf{A}$, let $d, e \in \mathbf{A}^{\mathbf{C}}$. Then:

$$d =_a e \Leftrightarrow d^{-1}(a) = e^{-1}(a) \text{ and } \sharp d = \sharp e$$

Let $a \in \mathbf{A}, B \subseteq \mathbf{A}$. We write $=_B := (\bigcup_{a \in B} =_a)^*$ and $=_{\cup B} := \bigcup_{a \in B} =_a$. Similarly for other equivalence relations.

Given a deal d, another deal e of the same size is *relevant* for the players in (that state of) the game, if e is $=_{\mathbf{A}}$-accessible from d, i.e. if e is it not publicly known to be 'irrelevant' in the common meaning of that word.

Definition 5 (Set of relevant deals)
Let $d \in \mathbf{A}^{\mathbf{C}}$. Then

$$D_d := [d]_{=_{\mathbf{A}}}$$

is the set of *relevant* deals given deal d.

Instead of $d' \in D_d$ we say that d' *is relevant*. This means that a player has to take d' in consideration when reasoning about an initial state of the game for deal d (see section 4, next). If there are more than two players and there is more than one card, then all deals in $D_{\sharp d}$ are relevant at the beginning of a knowledge game for an actual deal d. In the proof of this proposition, we use the following notion of *transposition* of a deal: let $d \in \mathbf{A}^{\mathbf{C}}$, then $d[c,c']$ is the deal such that $d[c,c'](c) = d(c')$, $d[c,c'](c') = d(c)$, and for all other cards $c'' \in \mathbf{C}$, $d[c,c'](c'') = d(c'')$. Note that $\sharp d[c,c'] = \sharp d$.

Proposition 1
If $|\mathbf{A}| > 2$, then for all $d \in \mathbf{A}^{\mathbf{C}}$, $D_d = D_{\sharp d}$.

Proof We may assume $\mathbf{C} \neq 0$ (otherwise, both D_d and $D_{\sharp d}$ are undefined). If there is one card only, all players know the deal of cards, because they can see who holds that card, and $D_d = D_{\sharp d} = \{d\}$. Now suppose there is more than one card. Let $e \in D_{\sharp d}$. Let $a_1 \in \mathbf{A}$. Let n be the number of cards that a_1 holds in d but doesn't hold in e, in other words: the number of differences between d and e. We now prove by induction on n that $d =_{\mathbf{A}} e$.

If $n = 0$ then $d^{-1}(a_1) = e^{-1}(a_1)$ (a_1 holds the same cards in d and e), thus $d =_{a_1} e$ and thus $d =_{\mathbf{A}} e$.

Suppose $d^{-1}(a_1)$ and $e^{-1}(a_1)$ differ in $n+1$ cards for player a_1. As there is more than one card, there are $c, c' \in \mathbf{C}$ such that $d(c) = a_1$, $e(c) \neq a_1$, $e(c') = a_1$, and $d(c') = a_2 \neq a_1$. As there are more than two players, there is a player $a_3 \neq a_1, a_2$. It holds that $d =_{a_3} d[c,c']$. As $d[c,c']$ differs in n cards from e for player a_1, by the induction hypothesis we can assume $d[c,c'] =_{\mathbf{A}} e$. From $d =_{a_3} d[c,c']$ and $d[c,c'] =_{\mathbf{A}} e$ follows $d =_{\mathbf{A}} e$. ■

A different way to express proposition 1 is to say that the equivalence relation $=_{\mathbf{A}}$ is the universal relation on $D_{\sharp d}$: $=_{\mathbf{A}} = D_{\sharp d} \times D_{\sharp d}$. One can even prove[2] that the maximum length of a path to link two arbitrary deals is at most three: $(=_{\cup \mathbf{A}})^3 = D_{\sharp d} \times D_{\sharp d}$. As $\langle D_{\sharp d}, (=_a)_{a \in \mathbf{A}} \rangle$ is nothing but a multiagent $S5$ *frame*, this property may help in characterizing it. We have not pursued that topic further. For the characterization of a different type of multiagent frame, see [Lom99, LvdMR00].

If there is only one player, he must necessarily hold all cards. If there are two players, each of them knows that the other player holds all other cards. Therefore, both players have full knowledge of the deal of cards. Only the actual deal of cards is relevant to them:

Fact 1
If there are only one or two players, $=_{\mathbf{A}}$ is the identity on $D_{\sharp d}$.

---

[2]Proof suggested by Gerard Renardel and by Josje Lodder. Let $d \in \mathbf{A}^{\mathbf{C}}$. Let $d_1 \neq d_2 \in D_{\sharp d}$. Let $a_1 \neq a_2 \neq a_3 \in \mathbf{A}$. Let $k = |d^{-1}(a_3)|$. If $k = 0$, then $d_1 =_{a_3} d_2$. Otherwise, choose k cards $c^1, ..., c^k$ from $\mathbf{C} \setminus (d_1^{-1}(a_1) \cup d_2^{-1}(a_2))$. Let d_3 be of size $\sharp d$ such that $d_3(c) = a_1 \Leftrightarrow d_1(c) = a_1$ and such that $d_3(c^1) = ... = d_3(c^k) = a_3$. Let d_4 be of size $\sharp d$ such that $d_4(c) = a_2 \Leftrightarrow d_2(c) = a_2$ and such that $d_4(c^1) = ... = d_4(c^k) = a_3$. Then $d_1 =_{a_1} d_3 =_{a_3} d_4 =_{a_2} d_2$. ■

4 State of the game

A model for the state of a knowledge game should contain all the information that the players have about the cards and about each other. Any game state is represented by a pointed multiagent $S5$ model. In the initial state of the game, players only know their own cards. In other game states, they may know more than that. We give some examples.

4.1 Initial state of a knowledge game

To represent the initial state of the game for the *actual deal* of cards $\mathsf{d} \in \mathbf{A}^{\mathbf{C}}$, we propose a pointed $S5$ model. Its worlds are deals, its domain is the set $D_{\sharp\mathsf{d}}$ of deals of the size of d, its point is the actual deal. The fourth letter of the alphabet in sans serif always denotes the actual deal. For each agent $a \in \mathbf{A}$, the accessibility relation $\sim_a$ between worlds is the equivalence relation $=_a$ as defined in definition 4: two worlds/deals are indistinguishable from each other for a, if they agree on his cards, and if in both deals all players hold the same number of cards. Before we can define a valuation on the worlds, we have to introduce atomic propositions: $\mathbf{P}$ is the set of $|\mathbf{C}| \cdot |\mathbf{A}|$ atomic propositions c_a corresponding to player $a \in \mathbf{A}$ holding card $c \in \mathbf{C}$. For any deal $e \in \mathbf{A}^{\mathbf{C}}$, V_e is the valuation such that $V_e(c_a) = 1 \Leftrightarrow e(c) = a$. We now define a global valuation $V : D_{\mathsf{d}} \to \mathbf{P} \to \{0,1\}$ that maps a deal e to such a (local) valuation V_e (with its argument, the deal e, as subscript). We sum it up in the following definition:

Definition 6 (Initial state of a knowledge game)
Let $\mathsf{d} \in \mathbf{A}^{\mathbf{C}}$, then the initial state of a game for actual deal of cards d is:

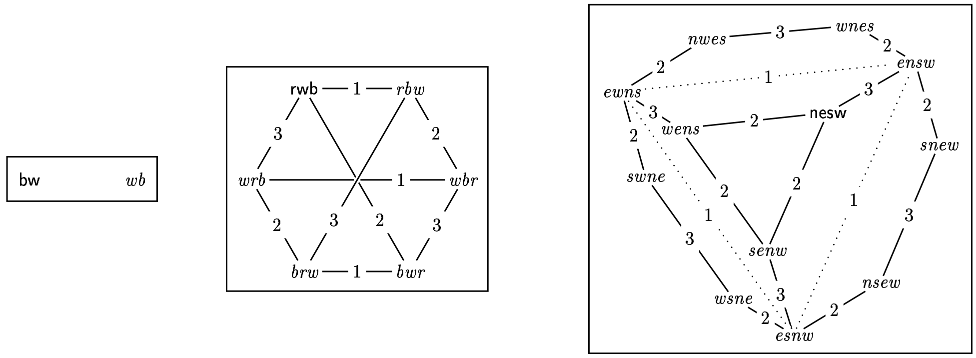
$$(\langle D_{\sharp\mathsf{d}}, \{\sim_a\}_{a\in\mathbf{A}}, V\rangle, \mathsf{d})$$

where:

$$\begin{array}{rrcl} \forall a \in \mathbf{A} : \forall d_1, d_2 \in D_{\sharp\mathsf{d}} : & d_1 \sim_a d_2 & \Leftrightarrow & d_1 =_a d_2 \\ \forall e \in D_{\sharp\mathsf{d}} : \forall c_a \in \mathbf{P} : & V_e(c_a) = 1 & \Leftrightarrow & e(c) = a \end{array}$$

Notation The model underlying an initial knowledge game state for actual deal d is written I_{d} (I for *i*nitial model), and thus the state itself is written $(I_{\mathsf{d}}, \mathsf{d})$. Instead of $(I_{\mathsf{d}}, \mathsf{d})$ we also write si_{d} (si for *s*tate and *i*nitial). Given the identification in definition 6, from now on we will generally write $\sim_a$ instead of $=_a$. We use the sans serif font both as a *graphical* device to distinguish the point of a model from its other worlds, and as a *textual* device mainly for the fourth letter of the alphabet to distinguish the actual deal from other deals: 'd' always means 'actual deal d', whereas in general the world 'w' that is the point of a model will be referred to as 'w' in the text.

Knowledge The relation between these models and the knowledge of the players is the following: in a given world w a player i *knows* something, if it



(I_{bw}, bw) (hexa, rwb) (tetra, $nesw$)

Figure 2: Examples of initial knowledge game states

holds in all (from w) $\sim_i$-accessible worlds. A player *can imagine* something (i.e. does not know the contrary), if it holds in some (from w) $\sim_i$-accessible world.

Figure 2 presents three initial knowledge game states. See also [vD00]. In the figures, we assume reflexive access for all worlds for all agents, as the relations $\sim_a$ are equivalence relations.

Example 2 (Letter)
The left picture in figure 2 represents the initial state $(I_{bw}, bw) = si_{bw}$ for the knowledge game for two players $1, 2$ and two cards b, w (black and white) with actual deal of cards $b|w$ (or bw). The point is in sans serif roman font: bw. Note that the other deal $w|b$ is not relevant given actual deal $b|w$.

From both worlds there is no access to other worlds. This corresponds to both players *knowing* what the (actual) deal of cards is. E.g. given deal bw, player 1 can only imagine that deal to be the case (he can only access that deal). Because he cannot imagine any other deal to be the case, player 1 *knows* that bw is the actual deal.

Example 3 (Hexa)
The middle picture in figure 2 represents the initial state si_{rwb} for the knowledge game for three players $1, 2, 3$ and three cards r, w, b with actual deal of cards $r|w|b$ (or rwb). The point is in sans serif roman font: rwb. We call the underlying model hexa. There are six different initial states of that game, corresponding to choosing a different deal as point in hexa. As the figure has the shape of a hexagon, it will now be clear why we call the corresponding game the hexa game.

Point rwb is the same for player 1 as rbw. In both deals he holds red. As he cannot access other worlds, he therefore *knows* that he holds red. From point rwb player 2 can access bwr. In bwr player 1 knows (similarly to the case rwb) that he holds blue. Therefore player 2 *can imagine* that player 1 knows that he holds blue. One more example: even though player 1 holds red, player 1 can imagine that player 2 can imagine that player 3 knows that he (3) holds red.

Example 4 (Tetra)
The right picture in figure 2 represents the initial state si_{nesw} for the knowledge game for three players $1, 2, 3$ and four cards n, e, s, w (north, east, south and west) with actual deal of cards $n|e|sw$ (or $nesw$). The point is in sans serif roman font: nesw. Access for agent 1 is only given in some typical cases. We call the underlying model tetra, because the figure has the shape of a semi-regular polyhedron called a truncated tetrahedron. To enhance the three-dimensional illusion of the picture we have not drawn access for agent 2 between worlds 'at the back', such as between $swne$ and $nwes$.

4.2 State of a knowledge game

After the cards have been dealt and everybody has seen his cards, players can learn about the cards of other players by means of game actions. Also other knowledge game states can be represented by pointed multiagent $S5$ models. We only require that agents *at least* know their own cards. In the next section, 5, we then define a game action; the execution of a game action induces a binary relation between such states.

Definition 7 (Knowledge game state)
A knowledge game state for deal of cards d is a pointed $S5$ model

$$(\langle W, \{\sim_a\}_{a \in \mathbf{A}}, V\rangle, v)$$

where $v \in W$, and $V_{\mathsf{d}} = V_v$, and:

$$\forall w \in W : \exists d' \in D_{\sharp\mathsf{d}} : V_w = V_{d'}$$

and for all $a \in \mathbf{A} : \forall w_1, w_2 \in W : \forall d_1, d_2 \in D_{\sharp\mathsf{d}}$:

$$(w_1 \sim_a w_2, V_{w_1} = V_{d_1}, V_{w_2} = V_{d_2}) \Rightarrow d_1 =_a d_2$$

Notation Unless confusion arises, we prefer to name worlds by the deals that atomically characterize them. So, worlds that are named by the same deal only differ in their access to other worlds. We then can continue to write d for the point of a state, instead of v. We often write s_v, or s_{d}, for a knowledge state for deal d with point v.

Initial knowledge game states are also knowledge game states. Use the convention to name worlds by deals. For knowledge game states we then have that $d_1 \sim_a d_2 \Rightarrow d_1 =_a d_2$. If additionally $d_1 =_a d_2 \Rightarrow d_1 \sim_a d_2$, it is an initial knowledge game state (more precisely: a model bisimilar to an initial knowledge game state).

As the game progresses, more and more deals of cards become irrelevant, in the sense that all players are known not to consider them any longer. This is captured by the following definition.

Definition 8
Let $s = (\langle W, \{\sim_a\}_{a \in \mathbf{A}}, V \rangle, v)$ be a knowledge game state for $\mathsf{d} \in \mathbf{A}^{\mathbf{C}}$, then:

$$D_s = \{d' \in D_{\mathsf{d}} \mid \exists w \in W : w \sim_{\mathbf{A}} v \text{ and } V_w = V_{d'}\}$$

Even though it is now clear what kind of mathematical objects knowledge game states are, this does not clarify what agents actually *know* and *don't know* in such a state of the game. For this, see [vD00].[3]

5 Game action

Given a knowledge game state, we now define game actions for that state. First we give an example of a game action and the knowledge state that we expect to result from its execution. Then, we speculate on a desirable format for game actions. Only after that, we present the definition of game actions. We conclude with applying the definition to the example action and with an overview of game actions in knowledge games.

Example 5 (1 shows red to 2)
In the initial state (hexa, rwb) of the hexa game, player 2 asks player 1 "please show me your card". The question is public: 3 hears it too. Player 1 answers this request (as in response (iii) related to example 1 on page 4) by handing his red card face down to player 2. Player 2 then looks at that card, after which he returns the card face down to player 1.

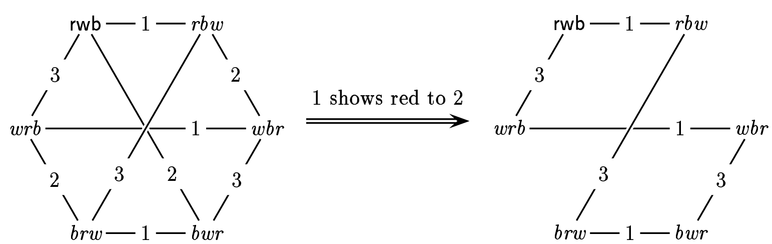


Figure 3: Player 1 shows (only) player 2 his red card

[3]In [vD00] we characterize the knowledge of the agents in initial knowledge game states, by describing these states in a multiagent epistemic logic with common knowledge operators. Because knowledge game states are finite models, their description can be computed with standard modal techniques, see [vB98, BM96]. The description of the model underlying the initial state of a knowledge game for deal $d \in \mathbf{A}^{\mathbf{C}}$ is equivalent to (the $\mathbf{S5_n}$ axioms plus) $\bigvee_{d' \in D_{\sharp d}} \delta_{d'}$ and $\bigwedge_{a \in \mathbf{A}} \bigwedge_{d' \in D_{\sharp d}} (K_a \delta^a_{d'} \leftrightarrow M_a \delta_{d'})$. Here, $\delta_{d'}$ is the atomic description of world (deal) d', and $\delta^a_{d'}$ is the part of that description about agent a, i.e. the conjunction of atoms c_a or their negations that describe the cards of a. The first expresses that each world is a deal of the size of d. The second expresses that for an agent deals are still possible if and only if they correspond to what he knows of his own cards.

The resulting state of the game is pictured on the right in figure 3. In the resulting game state, it should hold that player 2 knows that player 1 has the red card. Indeed, this is the case: as rwb is the only accessible world for player 2, he knows that rwb is the actual deal of cards. It follows that he knows that 1 holds red (and that 3 holds blue). Note that players 1 and 3 still don't know the deal of cards, as other worlds remain accessible for both 1 and 3. Compared to the initial game state, they still have learnt something: now they know, e.g., that player 2 knows the actual deal of cards! E.g., for player 1 worlds rwb and rbw are accessible. In both worlds it holds that 2 knows the deal of cards (see before). Therefore 1 knows that 2 knows the deal of cards.

Before we can compute game states resulting from game actions, we have to define what an action 'is'. In [vD00] we define a logical language to describe such actions. Here, we will investigate game actions from a purely semantical point of view: juggling with sets of deals, so to speak. So what is happening here? Player 2 knows that player 1 holds a card. In other words, 2 is aware of the partition of hexa induced by the equivalence relation $\sim_1$. By asking player 1 for his card, player 2 is presenting to player 1 the three different equivalence classes of $\sim_1$, for 1 to choose from: $\{rwb, rbw\}$, $\{wrb, wbr\}$, and $\{brw, bwr\}$. Player 1 hasn't much to choose, in this case, and has to answer by affirming that his information state (i.e. the set of worlds that he considers to be possible, see [vD00]) is $\{rwb, rbw\}$, which corresponds to 1 holding the red card. Player 3 does not receive the answer in the detail in which player 2 gets it. E.g. 3 cannot distinguish the answer red from the answer white.

A game action is a question with an answer to it We suggest that, just as in the action of example 5 where 1 shows red to 2, all game actions in knowledge games are the combination of a question with an answer to it. Another parameter of crucial importance is what other players perceive of the answer to the question, just as in the action where 1 shows red to 2. We call that the *publicity* of the game action.

Definition 9 (Game action)
Let $s = (\langle W, \{\sim_a\}_{a\in\mathbf{A}}, V\rangle, \mathsf{d})$ be a knowledge game state. A *game action* μ for state s is a quintuple

$$\mu = \langle q, Q, r, R, \mathsf{pub}\rangle$$

where $q, r \in \mathbf{A}$, Q is a covering of W that is coarser than $\sim_r$, $R \in Q$, and pub is a function from agents $a \in \mathbf{A}$ to equivalence relations pub_a on Q, and pub_r is the identity '=' on Q.[4]

In the definition, q is the requesting player, Q is the Question, r is the respondent, R is the answer or Response, and pub is the 'publicity': how and

[4] Q *covers* W if Q is a nonempty set of subsets of W such that $W = \bigcup Q$. Q is *coarser* than $\sim_r$ if $\forall x, x' : x \sim_r x' \to Q(x, x')$, where Q is seen as an equivalence relation inducing the partition Q. The result of both constraints is that all members of Q are the union of some equivalence classes of $\sim_r$.

what the respondent r makes public to other players of his answer to q. We can think of the elements $R_1, ..., R_{|Q|}$ of Q as 'possible answers' or 'alternative answers'. Obviously, r is informed about his own behaviour: the respondent can distinguish all possible answers from each other. In other words, the equivalence relation pub_r on the set of alternatives Q is the identity '='. We do not assume that the requesting player q is also fully informed, corresponding to $\mathsf{pub}_q =$ '=', although this is often a reasonable assumption.

Player 0 (the table), if there is one, cannot ask questions. Also it is only allowed to respond in certain reactive ways, and not proactively. What ways is determined by pub. You can draw a card from a stack on the table, and let the table respond to the request "show me one of your cards" in that way. You cannot ask the table "do you have the red or white card" and get "no" as an answer, or have it 'decide' which one of these two cards to show you.

Definition 10 (Executable game action)
Let $s = (\langle W, \{\sim_a\}_{a \in \mathbf{A}}, V \rangle, w)$ and $\mu = \langle q, Q, r, R, \mathsf{pub} \rangle$. Game action μ is *executable* in knowledge game state s if the answer R contains actual world w:

$$\mu \text{ is executable in } s \Leftrightarrow w \in R$$

In simpler words: if r answers the question truthfully. We may also say: if r's information state, i.e. $[w]_{\sim_r}$, is contained in the answer. In section 5.2, defintions 9 and 10 will be applied to example 5.

5.1 Publicity

Because the concept of 'publicity' is central to our approach, we give some motivations for it.

Knowledge games are all about getting information. The way to obtain information is to ask questions, with the expectation of getting certain answers, or to observe others asking and responding. We are only interested in (perfect) information: what other players know, what cards they hold; and not in strategic information: how likely it is they will ask a certain question, etc. Given this restriction, we can state that obtaining information (from questions and answers) is learning about the information state of the respondent. This 'learning' is not just individual but on the level of *subgroups* that gain *common knowledge* about the information state of the respondent.

There is a smallest nonempty subgroup $Br \subset \mathbf{A}$ – the broadcast unit so to speak – that receives the answer R to q.[5] Obviously $r \in Br$. Often, $q \in Br$. If the broadcast unit $Br = \mathbf{A}$, then the response R is publicly learnt. Otherwise,

[5] Strictly speaking, we mix up syntax and semantics here: one doesn't learn R but one learns a proposition φ_R with interpretation R. As our models are finite we can safely assume that such a proposition exists. See [vD00].

the broadcast unit Br is contained in at least one larger subgroup $B' \subseteq \mathbf{A}$. For the players in B' that are not in Br, that group Br learns answer R might be just one of several alternatives. For all they know, Br learns an alternative $R' \in Q$ that differs from response R. Or instead of Br learning R, a different subgroup $B'' \subset B'$ learns R. Even then, $r \in B''$, because the respondent r controls the publicity. Every such subgroup B' that is smaller than $\mathbf{A}$, is again contained in a larger one for which analogous restrictions hold. At some stage the entire group of agents $\mathbf{A}$ learn something: every action must have a public part.

If an action had no public part, some agents would learn nothing and would therefore think that nothing happened. In that case they have false knowledge of the state of the game: they 'know' that nothing happened, they 'know' that the broadcast unit Br has not learnt R, etc. As they do not consider the actual state of the game to be possible, the resulting pointed multiagent modal model is not reflexive and therefore not an $S5$ state, so certainly not a knowledge game state.

We only require that learning subgroups contain the respondent r, at whatever level of the transmission. We might additionally have required that subgroups also contain q, the requesting player. We haven't done that, because we also want to model actions such as: 'the respondent r showed a card to the requesting player q and his other card to player a'.

This may seem rather complex, but the very simple way to fulfill these constraints is to define for each agent a an equivalence relation pub_a on the set of alternative answers of a game action, as in definition 9. Having done that, for each subgroup $B \in \mathbf{A}$ we can, if so desired, compute $\mathsf{pub}_B = (\bigcup_{a \in B} \mathsf{pub}_a)^*$. *The equivalence class of* pub_B *that contains the answer* Q *stands for what subgroup* B *learns in that game action.*

5.2 Examples

Example 6 (1 shows red to 2, continued)
In example 5 we described the game action of 1 showing red to 2. This corresponds to the following game action:

$$\langle 2, \{\{rwb, rbw\}, \{wrb, wbr\}, \{brw, bwr\}\}, 1, \{rwb, rbw\}, \mathsf{show}\rangle$$

Player 2 asks the question. The question is $\{\{rwb, rbw\}, \{wrb, wbr\}, \{brw, bwr\}\}$, i.e. the three equivalence classes of $\sim_1$. Player 1 answers the question. The answer is $\{rwb, rbw\}$. This corresponds to 1 showing the red card. The publicity show is defined as follows: show_1 and show_2 are the identity '$=$' on the question, and show_3 is the universal relation U on the question.

According to definition 10, the actual deal of cards rwb should be contained in the given answer $\{rwb, rbw\}$. This is indeed the case. Therefore this game action is executable in initial state (hexa, rwb).

For player 3, the action where 1 shows red is indistinguishable from the action where 1 shows white. Because 3 holds blue himself, he does not consider that 1 actually shows blue, so that the action where 1 shows blue *can* be distinghuished by player 3 from the action where 1 shows red. However, because 1 doesn't know that 3 holds blue, 1 can imagine that 3 can imagine that 1 shows blue, or in other words: at some point or other, all three action alternatives have to be taken into account: 3 is not publicly known to be able to distinguish between the alternatives.

The observations on publicity in subsection 5.1 are mirrored by the computations we can do on publicity show in example 6. Using the equivalences $\mathsf{show}_1 = \mathsf{show}_2 =$ '=' and $\mathsf{show}_3 = U$, we can compute for every subgroup of the public $\{1, 2, 3\}$ what that subgroup has learnt. E.g. show_{12} is also the identity, whereas show_{13} is, again, the universal relation: 1 and 3 do not 'share' that the action of showing blue can be eliminated, as pointed out in the previous paragraph. Also, show_{123} is the universal relation: it is not publicly known which action has been taking place.

The next example illustrates why the alternatives are only required to *cover* the domain, and not to *partition* it. If the alternatives overlap, the responding player may *choose* from the alternatives that contain his information state:

Example 7
Assume initial knowledge game state (hexa, rwb). Consider the following action: Player 2 (publicly) asks player 1 "please tell me a card that you do not hold". As a response player 1 whispers in 2's ear "I do not hold white". Player 3 cannot hear the answer, although he knows that an answer has been given. This game action is described as follows; abbreviate $\{wrb, wbr, brw, bwr\}$ as R_{notred}, $\{rwb, rbw, brw, bwr\}$ as R_{notwhite}, and $\{rwb, rbw, wrb, wbr\}$ as R_{notblue}:

$$\langle 2, \{R_{\mathsf{notred}}, R_{\mathsf{notwhite}}, R_{\mathsf{notblue}}\}, 1, R_{\mathsf{notwhite}}, \mathsf{show}\rangle$$

Again, show_1 and show_2 are the identity and show_3 is the universal relation on the question. The alternative questions indeed cover the domain. E.g., $R_{\mathsf{notwhite}} = \{rwb, rbw, brw, bwr\}$ is the union of the two classes $\{rwb, rbw\}$ and $\{brw, bwr\}$ of $\sim_1$. The alternatives also overlap, e.g. R_{notwhite} and R_{notblue}: in the given state, player 1 could also have answered that he doesn't have blue.

In a knowledge game state where a player holds more than one card, he can choose between cards to show, given a request:

Example 8
Consider the state on the right in figure 2 on page 9, where player 3 holds the south and the west card. If player 2 asks player 3 for a card, player 3 may choose between showing south and showing west.

In [vD00], we introduce a multiagent dynamic epistemic language for describing game actions.[6]

5.3 Publicity functions and game actions in knowledge games

Only the following sorts of action occur in knowledge games: showing a card, not showing a card, winning, and not winning:

Definition 11 (Legal game actions in knowledge games)
Let $s = (\langle W, \{\sim_a\}_{a\in\mathbf{A}}, V\rangle, v)$ be a knowledge game state for a deal of cards $d \in \mathbf{A}^{\mathbf{C}}$. The sorts of action show, noshow, win and nowin are defined as follows; the middle column presents the abbreviated notations that we often use for them:

sort	*name*		*definition*
show	$\mathsf{show}^{q,\{c^1,\ldots,c^t\}}_{r,c^j}$	=	$\langle q, \{R_{c^1_r}, ..., R_{c^t_r}, Comp\}, r, R_{c^j_r}, \mathsf{show}\rangle$
noshow	$\mathsf{noshow}^{q,\{c^1,\ldots,c^t\}}_{r}$	=	$\langle q, \{R_{c^1_r}, ..., R_{c^t_r}, Comp\}, r, Comp, \mathsf{show}\rangle$
win	win^q	=	$\langle q, \{Win, W \setminus Win\}, q, Win, \mathsf{id}\rangle$
nowin	nowin^q	=	$\langle q, \{Win, W \setminus Win\}, q, W \setminus Win, \mathsf{id}\rangle$

We use the following abbreviations in the definition: $R_{c^i_a}$ stands for the union of equivalence classes of $\sim_a$ where a holds card $c^i \in \mathbf{C}$; $Comp$ stands for $W \setminus \cup_{i=1}^t R_{c^i_r}$, the complement of the union of all alternatives that correspond to r showing card c^i; Win stands for $\cup_{i=1}^t R_i$, the union of all equivalence classes R_i of $\sim_q$ where q can win. Publicity id maps each agent $a \in \mathbf{A}$ to the identity on the question, i.e. $\mathsf{id}_a =$ '='. Publicity show is defined as follows: show_q and show_r are the identity on the question, and for all other agents a, show_a is: *universal* on the question minus $Comp$, and the *identity* on that complement: $\mathsf{show}_a(Comp, Comp)$.

show and noshow In a show action player q asks player r to show him one of t cards $c^1, ..., c^t$, and r responds by showing (only) q that card. In a noshow action player q asks player r to show him one of t cards $c^1, ..., c^t$, and r responds by saying "no, I don't have any of those". Given some current state of the game, it can be that one or more of the R_{c^i} are empty. Also, if $t = |\mathbf{C}|$, the set $Comp$ is empty. In that case, we assume that they do not occur in the question: the question may contain only nonempty alternatives.

win and nowin The actions of winning and not winning are public announcements. In our game action format an announcement is a (public) question to

[6]For example, the *game action* of 1 showing red to 2 is described by the *knowledge action* $L_{123}(!L_{12}?r_1 \cup L_{12}?w_1 \cup L_{12}?b_1)$. Informally, we can read this expression as follows: 1 and 2 learn that 1 holds red, and 1, 2 and 3 learn that either 1 and 2 learn that 1 holds red, or that 1 and 2 learn that 1 holds white, or that 1 and 2 learn that 1 holds blue. The game action examples from this section are described as knowledge actions and are treated in more detail.

oneself that is publicly answered. If winning is knowing the actual deal of cards, Win is the union of all equivalence classes of $\sim_q$ that are characterized by a single deal of cards. If there are none, $Win = \emptyset$ and we assume that the question consists of W only: the question may contain only nonempty alternatives. Of course, $W \setminus Win$ corresponds to the union of all the equivalence classes where q cannot win.

The publicity functions show and id are the only common publicity functions that we have encountered.

5.4 Questions can always be answered

To illustrate how general the game action format is, we give two more examples.

Example 9 (What you don't know, is trivial)
Consider the state (hexa, rwb). Suppose player 2 asks player 1 "which card do I have?". Player 1 now answers "I don't know".

This appears to be a question that cannot be answered and therefore doesn't fit our game action format. It turns out that we can rephrase it to fit the format:
A question covers the domain of the game state, and is coarser than the partition induced on it by the equivalence relation for the responding player, in this case player 1. How does the question "which card do I have?" cover the domain of hexa? The informative answers to "which card do I have" are "red", "white" and "blue". Such an answer should contain the information state of 1, or, in other words, 1 should know that 2 holds that card. We first investigate the alternative corresponding to the answer "red". In no equivalence class of $\sim_1$, or union of classes, does 1 know that 2 holds red. Therefore, to this alternative corresponds the empty set $\emptyset$. Same for "white" and "blue". To these 'alternatives' we add the complement of their union, in this case the entire domain of hexa: $\mathcal{D}(\mathsf{hexa})$. The answer "I don't know" corresponds to choosing that complement. Therefore this game action is represented by:

$$\mu = \langle 2, \{\mathcal{D}(\mathsf{hexa})\}, 1, \mathcal{D}(\mathsf{hexa}), \mathsf{id} \rangle$$

This is a trivial game action, where player 1 asserts that his information state is contained in the domain of the state of the game. This was already known to all players, so that there is nothing to be gained from executing this game action: the resulting game state is, again, (hexa, rwb). In the initial state of the game it is senseless to ask others about your own cards, and everybody knows that.

A question is commonly considered 'trivial', if the person asking knows in advance the answer to his question. We suggest that a game action is trivial in a technical sense, if that is the case for all players.

Definition 12 (Trivial action / question)
A game action $\langle q, Q, r, R, \mathsf{pub} \rangle$ for state s_w is trivial, when for all $a \in \mathbf{A}$: $[w]_{\sim_a} \subseteq R$ and $\forall R' \neq R \in Q : w \notin R'$.

Without the constraint expressing that w is only in R, the responding player might have chosen a possibly informative (and therefore non-trivial) answer $R' \neq R$. Example 9 was an illustration of this definition.

We have thought of several other sorts of game action, apart from the five introduced in this and in the previous subsection, but we do not pursue this topic further.

We have modelled questions and answers in *games*. The logical modelling of questions and answers is studied more generally by Jeroen Groenendijk in [Gro99], 'The Logic of Interrogation'. In his approach, a question induces a *partition* on a *part* of the domain of a state. In our approach, questions induce a *covering* of the *entire* domain. Because questions are not partial on the domain, they can always be answered. Because possible answers can overlap, the respondent may choose between alternative answers to a question. Also because of that, the model resulting from execution of a game action, i.e. a question/answer combination, may be more complex than the model in which the question was posed. First, we have to define the construction of that resulting model.

6 Computing the next state of the game

We know what knowledge game states are and how to model game actions. We still have to define what game state results from executing a game action. We do not need all the parameters of a game action for that: who asked the question and who responded to it, is irrelevant for computing the information changes. Stripped from these two parameters, what remains is a multiagent $S5$ frame that we call a game action frame.

Definition 13 (Game action frame)
Let $\mu = \langle q, Q, r, R, \mathsf{pub} \rangle$ be a game action. To this game action corresponds the pointed multiagent $S5$ frame:

$$\mu^- = (\langle Q, \mathsf{pub} \rangle, R)$$

Note that the frame may consist of disconnected parts, both when the publicity of the action is id as when it is show.

We can now define how to compute the next state of a game from a given knowledge game state and a game action. A knowledge game state is represented by a pointed $S5$ model. A game action is represented by a pointed $S5$ frame. The computation of the next game state from the current state and an action, or in other words the execution of that action in that state, can be seen as *multiplying* the pointed $S5$ model for that state with the pointed $S5$ frame for that action: it resembles the computation of a direct product (see [vD00]).

Definition 14 (Executing a game action in a knowledge game state) Let $s = (\langle W, \{\sim_a\}_{a \in \mathbf{A}}, V\rangle, v)$ be a knowledge game state for actual deal d and let $\mu = \langle q, Q, r, R, \mathsf{pub}\rangle$ be a game action executable in s. The knowledge game state $s \otimes \mu$ resulting from executing μ in s is defined as follows:

$$s \otimes \mu := (\langle W', (\sim'_a)_{a \in \mathbf{A}}, V'\rangle, (v, R))$$

$$\begin{array}{rcl} & \text{where:} & \\ W' & = & \{(w, R') \in W \times Q \mid w \in R'\} \\ \text{and } \forall a \in \mathbf{A} : \forall w, w' \in W : \forall R', R'' \in Q : & & \\ (w, R') \sim'_a (w', R'') & \Leftrightarrow & w \sim_a w' \text{ and } \mathsf{pub}_a(R', R'') \\ V'_{(w,R')} & = & V_w \end{array}$$

The general idea of this construction is, that the next state of the game consists of all pairs (w, R') such that R' 'could also have been' the given answer and w 'could also have been' the (point of the) current state, plus access appropriately defined. The computation of $s \otimes \mu$ does not depend on the roles of the player asking the question and the player responding to it. Therefore, instead of $s \otimes \mu$ we may also write $s \otimes \mu^-$, where μ^- is the pointed frame corresponding to game action μ. We still have to prove that the resulting model $s \otimes \mu$ is a knowledge game state. This is indeed the case:

Proposition 2 ($s \otimes \mu$ is a knowledge game state) Let $s = (\langle W, \{\sim_a\}_{a \in \mathbf{A}}, V\rangle, v)$ be a knowledge game state for deal d, and let $\mu = \langle q, Q, r, R, \mathsf{pub}\rangle$ be a game action executable in s. Then $s \otimes \mu$ is a knowledge game state for deal d.

Proof We check the requirements from definition 7:

- $V'_{(v,R)} = V_{\mathsf{d}}$:
  This follows from $V'_{(v,R)} = V_v$ and $V_v = V_{\mathsf{d}}$.
- every world in W' is characterized by a(n) (initially) relevant deal:
  This follows from $V'_{(w,R')} = V_w = V_{d'}$ for some $d' \in D_s \subseteq D_{\mathsf{d}}$.
- for all $a \in \mathbf{A}$, $\sim'_a$ is an equivalence relation, which is obvious, such that:
- if players cannot distinguish between two worlds, they hold the same cards in (the deals that characterize) those worlds:
  If $(w, R') \sim'_a (w', R'')$, then $w \sim_a w'$, so a holds the same cards in w and w' and therefore also in (w, R') and (w', R''). ■

In general, it does not hold that the product of two connected structures is connected. However, when executing game actions in game states, this property is preserved:

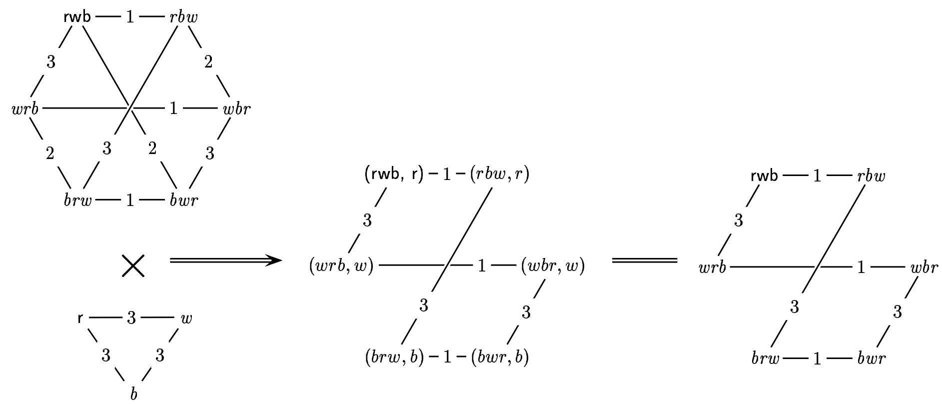


Figure 4: Executing action $\mathsf{show}^{2,-}_{1,r}$ in state $(\mathsf{hexa}, \mathsf{rwb})$

Proposition 3 (Preservation of connectedness)
Let $s = (\langle W, \{\sim_a\}_{a\in\mathbf{A}}, V\rangle, v)$ be a knowledge game state for deal d, let $\mu = \langle q, Q, r, R, \mathsf{pub}\rangle$ be a game action executable in s. Write $s \otimes \mu$ as above.

If $[v]_{\sim_\mathbf{A}} = W$ and $[R]_{\mathsf{pub}_\mathbf{A}} = Q$, then $[(v, Q)]_{\sim'_\mathbf{A}} = W'$.

Proof Let $(w, R'), (w', R'') \in W'$. From $w, w' \in W$ follows $w \sim_\mathbf{A} w'$. From $R', R'' \in Q$ follows $\mathsf{pub}_\mathbf{A}(R', R'')$. Because $w \sim_\mathbf{A} w'$ there is a finite chain $w \sim_{a_1} \ldots \sim_{a_n} w'$. Also, because for all $a \in \mathbf{A}$, pub_a is an equivalence relation and therefore reflexive, we have (writing infix) $R'\ \mathsf{pub}_{a_1} \ldots \mathsf{pub}_{a_n}\ R'$; therefore, using the definition of $\sim'$: $(w, R') \sim'_{a_1} \ldots \sim'_{a_n} (w', R')$, and thus $(w, R') \sim'_\mathbf{A} (w', R')$. Similarly for $(w', R') \sim'_\mathbf{A} (w', R'')$. We now have $(w, R') \sim'_\mathbf{A} (w', R') \sim'_\mathbf{A} (w', R'')$, and therefore, as $\sim'_\mathbf{A}$ is an equivalence relation, $(w, R') \sim'_\mathbf{A} (w', R'')$. ■

Example 10 (Executing $\mathsf{show}^{2,-}_{1,r}$)
We apply definition 14 to the knowledge game state $(\mathsf{hexa}, \mathsf{rwb})$ and the game action $\langle 2, \{\{rwb, rbw\}, \{wrb, wbr\}, \{brw, bwr\}\}, 1, \{rwb, rbw\}, \mathsf{show}\rangle$ also abbreviated as $\mathsf{show}^{2,-}_{1,r}$. Indeed the computations then result in the knowledge game state on the right in figure 3 on page 11. In figure 4 we visualize the construction. In the figure, we abbreviate $\{rwb, rbw\}$ as r, $\{wrb, wbr\}$ as w, and $\{brw, bwr\}$ as b. E.g. in the middle figure we have that $(rbw, r) \sim_3 (brw, b)$ because $rbw \sim_3 brw$ and $\mathsf{show}_3(r, b)$. In the figure on the extreme right, we follow the convention that we name worlds by the deals that atomically characterize them.[7]

[7] We mentioned in subsection 5.2 that the knowledge action $L_{123}(!L_{12}?r_1 \cup L_{12}?w_1 \cup L_{12}?b_1)$ corresponds to the game action $\mathsf{show}^{2,-}_{1,r}$. We can now explain what that correspondence is:

The knowledge action $L_{123}(!L_{12}?r_1 \cup L_{12}?w_1 \cup L_{12}?b_1)$ is interpreted as a state transformer, i.e. as a relation between states. The pointed frame for the game action $\mathsf{show}^{2,-}_{1,r}$ is a semantic object. However, its execution induces exactly the same relation between knowledge states as that of the corresponding knowledge action. See [vD00].

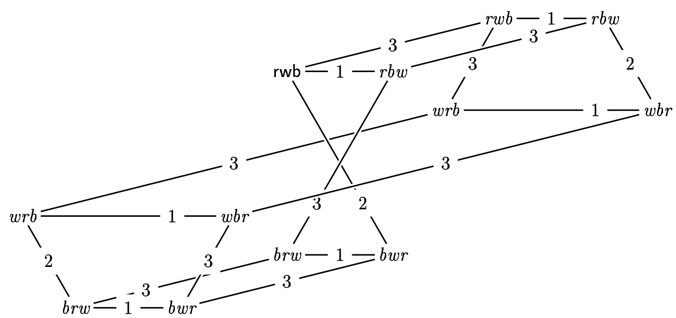


Figure 5: Player 1 has whispered in 2's ear: "I do not hold white".

Example 11
If a player can *choose* how to respond to a question, the next state of the game may be more complex, as measured in the number of nonsimilar worlds, than the current state of the game. Figure 5 pictures the result of executing in state (hexa, rwb) the game action where player 1 whispers in 2's ear "I do not hold white", given a request for a card that he does not hold; see example 7. The resulting state consists of 12 worlds. We assume transitivity in the figure. Note that in the point rwb (sans serif) of that state, neither 2 nor 3 knows that 1 holds red: this is, because 2 can access a world bwr where 1 does not hold red, and because 3 can access a world wrb where 1 does not hold red (by way of the world rwb in italics, 'at the back'). The other world rwb (italics) of that state corresponds to the answer "I do not hold blue". In that world player 2 now indeed knows that 1 holds red.

In particular, note that in the last example, 11, execution of the action does not result in the refinement of partitions on the given set of states, but in an increased number of states, as characterized both by the description of facts about the world, and by beliefs (in our case: knowledge) of the players about each other. The same holds for the action in example 8 on page 15.

7 Playing knowledge games

In the previous sections we have defined what a state of the game is, what an action in the game is, and how to compute the next state of the game. We have not actually played knowledge games. It will be clear that only with all this groundwork covered we can start to think about optimal strategies for playing knowledge games. To investigate this is relevant for game theory, because playing knowledge games is nothing but proceduralized information exchange in groups of competitors, where the value of questions and answers depends on their content and on the group members that receive this information.

The notion of an action as a semantic object is similar to that in [Bal99]. For the relation to Baltag, see also [vD00].

First, we have to define the *rules* of the game. Then, we can describe strategies. Only then, we can compute optimal strategies. In this section, we will touch upon these different topics.

Rules for playing knowledge games The game actions that occur in knowledge games are of the sort: show, noshow, win and nowin. A player can ask another player for one of some cards, after which the other player replies by saying that he doesn't have them, or by showing to the requesting player one of the requested cards. And a player can declare to have won the game, or, by finishing his move, 'declare' that he cannot win yet.

The order protocol restricts the order of actions in a game. These restrictions are computed from the history of requesting players, i.e. from the first argument of game actions. A play of the game consists of an alternating sequence of either show or noshow actions followed by nowin actions, where the last of the show or noshow actions is followed by a win action, the final action in the play. This means that, after a show or noshow action, we always allow the requesting player to announce that he has won. We could have chosen a different protocol, where he also is allowed to *guess*, see below.

It is randomly determined who starts the game, i.e. who asks the first question. The table is, naturally, *not* allowed to ask a question. Drawing a card from the table is not permitted as an action.

Guessing We only allow players to announce that they *know* the deal of cards (or to announce knowledge of another winning condition). Instead, we may allow players to *guess* the deal of cards. This is just as in the real Cluedo game. A player may only guess once during the game; if his guess is wrong, he lost. We can define a new kind of knowledge game: a player may *either* ask for a card, in which case a show or noshow action results, *or* he may guess the deal of cards, in which case a win or nowin action results.[8] He may not do both during his turn, as we allowed before. Therefore it is no longer the case, that after a show or noshow action, a player performs a nowin action, unless he wins: now a player may know, but he is only allowed to announce that knowledge in his next turn. In the 'guessing game' a show or noshow action is always *preceded* by a nowin action, instead of *almost* always followed by it. This suggests that we can redefine the game actions show and noshow by adding an extra constraint on the question.

Winning the game In the example of three players and three cards, winning was publicly announcing knowledge of the deal of cards. We can define weaker criteria for knowledge required to win, e.g. knowledge of the cards of one particular player. An example is knowledge of the cards on the table, as in Cluedo.

Game theory Knowledge games are competitive games of imperfect infor-

[8] Suggested by Ariel Rubinstein, personal communication

mation, where the only final outcomes are that players can win or lose. (See e.g. [OR94, Bin92] for a general introduction.) The value of such a game is the probability for the starting player to win, given that players follow optimal strategies. These are mixed strategies. For very simple knowledge games we can draw the entire game tree and by backwards induction compute the value of the game. It will be obvious that the value of a hexa game is 1: the beginner can always win in the first move. A slightly more complex knowledge game is that for two players and five cards, where each player holds two cards and one card is lying on the table. Winning is announcing knowledge of the card lying on the table. Assume that players are not allowed to ask the same question twice, so that the game tree is finite. Both when a player is only allowed to ask for one card, and when a player is only allowed to ask for one of two cards, the value of the game is $\frac{7}{9}$. Does the first player always have the highest probability to win a knowledge game? What is the value of Cluedo? In general, we have not answered these questions. The individual preference relation for a player on the different questions that he can ask, depends on the answers that he expects, on the probability distribution of these answers, and on how they refine the partitions for all players. This is quite hard to compute, if possible at all, see [Koo00]. We intend to investigate this in future research.

8 Conclusion

We have defined the concepts of knowledge game, deal of cards, knowledge game state, game action, and action execution. Questions and answers in games can be modelled as game actions. We now can describe in mathematical detail actual card games and card requests and responses in those games. Only given this precise definition of game states and game actions, can we start to think about optimal strategies for playing such games. Our results are relevant to the analysis of communication in groups.

References

[Bal99] A. Baltag. A logic of epistemic actions. Manuscript, 1999.

[Bin92] K. Binmore. *Fun and Games.* D.C. Heath, Lexington MA, 1992.

[BM96] J. Barwise and L.S. Moss. *Vicious Circles.* CSLI Publications, Stanford CA, 1996.

[FHMV95] R. Fagin, J.Y. Halpern, Y. Moses, and M.Y. Vardi. *Reasoning about Knowledge.* MIT Press, Cambridge MA, 1995.

[Ger99] J.D. Gerbrandy. *Bisimulations on Planet Kripke.* PhD thesis, University of Amsterdam, 1999. ILLC Dissertation Series DS-1999-01.

[Gro99] J. Groenendijk. The logic of interrogation. In T. Matthews and D.L. Strolovitch, editors, *Proceedings of the Ninth Conference on Semantics and Linguistic Theory*. CLC Publications, 1999.

[Koo00] B.P. Kooi. Master(s)mind(s). Manuscript, 2000.

[Lom99] A.R. Lomuscio. *Knowledge Sharing among Ideal Agents*. PhD thesis, University of Birmingham, Birmingham, UK, 1999.

[LvdMR00] A.R. Lomuscio, R. van der Meyden, and M.D. Ryan. Knowledge in multiagent systems: Initial configurations and broadcast. *ACM Transactions on Computational Logic*, 2000. to appear.

[OR94] M.J. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, Cambridge MA, 1994.

[vB98] J.F.A.K. van Benthem. Dynamic odds and ends. Technical report, ILLC, University of Amsterdam, 1998. Report ML-1998-08.

[vD00] H.P. van Ditmarsch. *Knowledge games*. PhD thesis, University of Groningen, 2000. ILLC Dissertation Series DS-2000-06.