Descriptions of game states

Hans van Ditmarsch, Wiebe van der Hoek, Barteld Kooi

Knowledge games are a fair playground to study properties of multiagent systems. In the initial state of a knowledge game a number of cards is dealt over a number of players. We describe this initial game state by a theory comprising facts about deals of cards and the agents' knowledge and ignorance of those facts. We also provide a Kripke ($S5$) model for such a game state and show that it is unique up to bisimulation. Therefore this model demonstrates all and only the relevant epistemic properties of the initial state.

1 Introduction

The interaction between logic and game theory is currently of interest to the scientific community. Well-known are game theoretical foundations for logical semantics (Hintikka and Sandu, 1989), and other applications of game theory in logic (Ehrenfeucht, 1961; van Benthem, 2000a). For applications of logic in game theory, we may mention the formalization in logical theories of game theoretical notions such as game trees, plays of a game, and equilibria (Bonanno, 1993; Kaneko and Nagashima, 1996). One issue of interest in this area are games where the information contained in a game state and the information change due to a game action may be rather complex, and therefore become objects of study in themselves. As a concrete example of such games we have defined knowledge games (van Ditmarsch, 2001): card games where a number of cards is distributed over a number of players (in the context of games, we prefer to call agents 'players'), and where moves consist of information exchange, such as showing cards to other players. Of particular interest are actions where information is simultaneously exchanged between subgroups of different size, e.g. when one player shows a card to another player, while the remaining players see that a card is being shown but not which card it is.

From a given knowledge game state and a game action that is executable in that state we can compute the next game state (van Ditmarsch, 2000). Therefore, we can compute any game state from an

Submitted for
Proceedings of LLC9
Ingrid van Loon, Grigori Mints, Reinhard Muskens (eds.)

initial knowledge game state and a game action sequence. This illustrates the need for a logical description of initial game states. Although it seems to be rather clear what the players know about *facts* in an initial game state, the precise amount of their ignorance of *other players' knowledge* is less transparent, as one easily overlooks game features. It is important to know what other players know, because in other game states this may help one to derive factual knowledge, which may help to win.

In this article we provide descriptions of initial knowledge game *states*. For the description of game *actions*, see (van Ditmarsch, 2000).

In section 2 we give an overview of relevant logical terminology. Readers familiar with epistemic logic may prefer to skip this section. See also (Meyer and van der Hoek, 1995; Fagin et al., 1995). In section 3 we present the theory Hexa that describes the model $Hexa$ of the initial game state of the knowledge game for three players each holding one card. This knowledge game exemplifies most of the features that we want to study. In section 4 we continue with the general case: the theory Kgames describes the initial game state of the knowledge game for an arbitrary deal of cards over players. In section 5 we describe the game state where the cards have been dealt but where players haven't picked up and looked at their cards.

2 Epistemic logic

Given (throughout the text) are a finite set A of agents and a set P of atoms.

Syntax of epistemic logic Multiagent epistemic logic $\mathcal{L}_A^P$ is the smallest set such that, if $p \in P, \varphi, \psi \in \mathcal{L}_A^P, a \in A, B \subseteq A$, then: $p, \neg\varphi, (\varphi \wedge \psi), K_a\varphi, C_B\varphi \in \mathcal{L}_A^P$. Formula $K_a\varphi$ means '*a knows φ*'. Formula $C_B\varphi$ means '*B know φ*' (or '*B commonly know φ*'). If $B = A$ we also say that '*B publicly know φ*'. We introduce the usual abbreviations $\varphi \vee \psi$, $\varphi \rightarrow \psi$, and $\varphi \leftrightarrow \psi$.

Epistemic structures An $S5_A$ *model* is a triple $\langle W, \{\sim_a\}_{a\in A}, V\rangle$ where W is the (nonempty) *domain*, for each $a \in A$, $\sim_a \subseteq W \times W$ is the *accessibility relation* for agent a, which is an *equivalence relation*, and the *valuation* V is a function that given a world assigns a truth value to each atom: $V : W \rightarrow P \rightarrow \{0,1\}$ (i.e. for each $w \in W$, $V_w : P \rightarrow \{0,1\}$). Instead of $w \in W$ we also write $w \in M$. The pair $\langle W, \{\sim_a\}_{a\in A}\rangle$ is an $S5_A$ *frame*. If $w \in M$, the pair (M, w) is an $S5_A$

state, where w is its *point* or *designated world*.

Define $\sim_B \;:= (\bigcup_{a\in B} \sim_a)^*$ (* stands for transitive and reflexive closure). An $S5_A^C$ model is an $S5_A$ model with accessibility relations $\sim_B$ added for all $B \subseteq A$. Every $S5_A^C$ model can also be seen as an $S5_A$ model. We generally write $S5$ instead of $S5_A$, and we always assume an arbitrary model to be an $S5$ model.

Semantics of epistemic logic Let $M = \langle W, \{\sim_a\}_{a\in A}, V\rangle$ be an $S5$ model, $w \in M$, $p \in P$, $a \in A$, $B \subseteq A$. Then: $M, w \models p \Leftrightarrow V_w(p) = 1$; $M, w \models \neg\varphi \Leftrightarrow M, w \not\models \varphi$; $M, w \models \varphi \wedge \psi \Leftrightarrow [\ M, w \models \varphi$ and $M, w \models \psi\ ]$; $M, w \models K_a\varphi \Leftrightarrow \forall w' \sim_a w : M, w' \models \varphi$; $M, w \models C_B\varphi \Leftrightarrow \forall w' \sim_B w : M, w' \models \varphi$. Derived notions are: $M \models \varphi \Leftrightarrow \forall w \in M : M, w \models \varphi$ (note that we have $M \models \varphi \Leftrightarrow M \models C_A\varphi$), and $M, w \models \Sigma \Leftrightarrow \forall \varphi \in \Sigma : M, w \models \varphi$.

We now define (local) logical consequence. We write $\models$ instead of, strictly, $\models_{S5_A}$. We define $\Sigma \models \varphi \Leftrightarrow \forall M, \forall w \in M : M, w \models \Sigma \Rightarrow M, w \models \varphi$.

Bisimulation Let $M = \langle W, \{\sim_a\}_{a\in A}, V\rangle$ and $N = \langle W'', \{\sim''_a\}_{a\in A}, V''\rangle$ be $S5$ models. A *bisimulation* between two states (M, x) and (N, y) is a nonempty binary relation $\mathfrak{R} \subseteq W \times W''$ such that $\mathfrak{R}(x, y)$ and:

Atoms: $\forall w \in M, \forall v \in N : \mathfrak{R}(w, v) \Rightarrow V_w = V''_v$;

Forth: $\forall a \in A, \forall w, w' \in M, \forall v \in N : \mathfrak{R}(w, v)$ and $w \sim_a w' \Rightarrow \exists v' \in N : v \sim''_a v'$ and $\mathfrak{R}(w', v')$;

Back: $\forall a \in A, \forall w \in M, \forall v, v' \in N : \mathfrak{R}(w, v)$ and $v \sim''_a v' \Rightarrow \exists w' \in M : w \sim_a w'$ and $\mathfrak{R}(w', v')$.

We say that (M, x) *is bisimilar to* (N, y) and we write $(M, x) \underline{\leftrightarrow} (N, y)$. If $\mathsf{domain}(\mathfrak{R}) = M$ and $\mathsf{range}(\mathfrak{R}) = N$, M *is bisimilar to* N, and we write $M \underline{\leftrightarrow} N$. It holds that $(M, w) \underline{\leftrightarrow} (N, v) \Rightarrow [\ \forall\varphi : M, w \models \varphi \Leftrightarrow N, v \models \varphi\ ]$, and that $M \underline{\leftrightarrow} N \Rightarrow [\ \forall\varphi : M \models \varphi \Leftrightarrow N \models \varphi\ ]$.

Description Let M be an $S5$ model, $w \in M$. The *atomic description* φ of world w (or of state M, w) is $\bigwedge_{p\in P} p^w$, where $p^w := p$ if $V_w(p) = 1$ and $p^w := \neg p$ if $V_w(p) = 0$. A *description of a model* M is a set of formulas (i.e. a *theory*) Σ such that $M \models \Sigma$ and $\forall M' : M' \models \Sigma \Rightarrow M' \underline{\leftrightarrow} M$. In other words: M is a model of Σ, and *only* M is a model of Σ. A *description of a state* (M, w) is a set Σ such that $M, w \models \Sigma$ and $\forall M', \forall w' \in M' : M', w' \models \Sigma \Rightarrow (M, w) \underline{\leftrightarrow} (M', w')$.[1] Finally, a *characteristic* φ of M is merely a formula that is valid in M: $\forall w \in M : M, w \models \varphi$.

[1] Occasionally we also use ‘describe’ in an informal sense, as in ‘atom c_a describes that player a holds card c’.

3 Description of Hexa

In this section, we present the theory Hexa, that describes the ($S5_{\{1,2,3\}}$) model $Hexa$. The state ($Hexa, rwb$) is the initial state of the knowledge game for three players each holding a card, where 1 holds red (r), 2 holds white (w), and 3 holds blue (b). In such games, moves consist of questions, such as "Do you have the white or the blue card?", and their answers, such as "No". We disregard all game aspects and merely focus on the logical description of the initial game state. Figure 1.1 pictures the model $Hexa$. In the figure, a node labelled cde represents the deal of cards where 1 holds c, 2 holds d and 3 holds e; an arc is labelled with a if player a cannot distinguish from each other the deals that are linked by it.

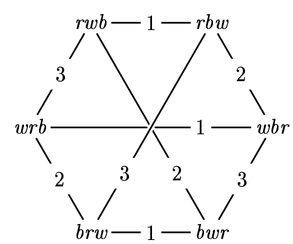

Figure 1.1: The model $Hexa$ for three players each holding a card

What information do the players have in this model, regardless of the actual deal of cards? They know how many cards there are, namely three. They know that the cards are all different, namely one red, one white and one blue. They know that each of them holds one card. Beyond that, if they hold a card, they know it, and if they don't hold a card, they also know that they do not hold it. All this is publicly known. They don't know anything else, and there seem to be two sides of that ignorance. First, a player doesn't know that another player holds a specific card. Second, with the exception of his own card, a player can imagine any card to be in possession of another player. All these constraints are satisfied by the theory Hexa.

Definition 1 (Hexa) *Theory* Hexa *consists of the formulas:*

$$\begin{array}{rcl} \mathsf{seeH} & := & \bigwedge_{a\in\{1,2,3\}} \bigwedge_{c\in\{r,w,b\}} (c_a \rightarrow K_a c_a) \\ \mathsf{dealsH} & := & \delta_{rwb} \vee \delta_{rbw} \vee \delta_{wrb} \vee \delta_{wbr} \vee \delta_{brw} \vee \delta_{bwr} \\ \mathsf{dontknowthatH} & := & \bigwedge_{a\neq b\in\{1,2,3\}} \bigwedge_{c\in\{r,w,b\}} \neg K_a c_b \end{array}$$

These formulas are the *constituents* of the theory. Atomic proposition c_a expresses that player a holds card c. SeeH expresses that every agent can see his own card. DealsH expresses that there are only six relevant deals, δ_{abc} is the atomic description of the world (deal) abc in $Hexa$, i.e. $a_1 \wedge \neg b_1 \wedge \neg c_1 \wedge \neg a_2 \wedge b_2 \wedge \neg c_2 \wedge \neg a_3 \wedge \neg b_3 \wedge c_3$. DontknowthatH expresses that players do not know the cards of other players. In section 3.1 we discuss various properties of agent knowledge in $Hexa$ that can be derived from the concise formulation in Hexa.

Theorem 1 *Theory* Hexa *describes model* $Hexa$.

We have to show that $Hexa \models \mathsf{Hexa}$, which is obvious, and that only $Hexa$ models Hexa. The last follows from the following lemma, that is a special case of lemma 6, to be presented and proven in section 4:

Lemma 2 *Suppose* $M \models \mathsf{Hexa}$. *Then* $M \underline{\leftrightarrow} Hexa$.

What does it mean that Hexa describes $Hexa$? Only in a technical sense is $Hexa$ the only model of Hexa: Hexa defines the bisimulation class of $Hexa$. But from this it follows that every non-theorem of Hexa is falsified in $Hexa$, and every theorem of Hexa is valid in $Hexa$. Differently said, if we *weaken* theory Hexa, e.g. by deleting formulas, it models structures of unintended game states, but if we *strengthen* the theory, it no longer models $Hexa$.

When we say that Hexa describes $Hexa$, we have implicitly quantified over all its worlds. This has to be made explicit when we describe one of its states: we now demand that Hexa is common knowledge in that state, and that its (unique) atomic description holds:

Corollary 3 *Theory* $C_{123}\mathsf{Hexa} + \delta_d$ *describes* $(Hexa, d)$.[2]

Corollary 4 *For all* $\mathcal{L}_{123}^{\{r_1, r_2, \cdots\}}$ *formulas* φ *and* ψ: $Hexa \models \varphi \Leftrightarrow Hexa \models C_{123}\varphi \Leftrightarrow C_{123}\mathsf{Hexa} \models \varphi \Leftrightarrow C_{123}\mathsf{Hexa} \models C_{123}\varphi$ *and* $Hexa, d \models \psi \Leftrightarrow C_{123}\mathsf{Hexa} + \delta_d \models \psi$.

Corollary 4 is a direct consequence of theorem 1. Note that $\mathsf{Hexa} \models \varphi \not\Leftrightarrow \mathsf{Hexa} \models C_{123}\varphi$, as e.g. $\mathsf{deals} \not\models C_{123}\mathsf{deals}$.

In a way, theorem 1 also shows that we have chosen the *right* model, and the *right* description, for the game state of three players each holding a card. From the viewpoint of knowledge specification, we were initially uncertain about *both* the properties of the agents in this game state *and* the $S5$ model representing it. But by linking these so tightly

[2] Where $C_{123}\mathsf{Hexa} + \delta_d = \{C_{123}\varphi \mid \varphi \in \mathsf{Hexa}\} \cup \{\delta_d\}$.

together in theorem 1 we have simultaneously validated them, so to speak.

3.1 Derived characteristics of Hexa

One can define various other characteristics of $Hexa$. We list a few:

Definition 2 (Other properties of players' knowledge)

$$\begin{array}{rcl} & & \text{players only see their own cards} \\ \mathsf{dontseeH} & := & \bigwedge_{a\in\{1,2,3\}} \bigwedge_{c\in\{r,w,b\}} (\neg c_a \rightarrow K_a \neg c_a) \\ & & \text{there is at most one card of each colour} \\ \mathsf{atmostH} & := & \bigwedge_{a\neq b\in\{1,2,3\}} \bigwedge_{c\in\{r,w,b\}} \neg(c_a \wedge c_b) \\ & & \text{there is at least one card per player} \\ \mathsf{atleastH} & := & \bigwedge_{a\in\{1,2,3\}} (r_a \vee w_a \vee b_a) \\ & & \text{players can imagine that others hold other cards} \\ \mathsf{dontknownotH} & := & \bigwedge_{a\neq b\in\{1,2,3\}} \bigwedge_{c\in\{r,w,b\}} (\neg c_a \rightarrow \neg K_a \neg c_b) \\ & & \text{players do not know that others hold other cards} \\ \mathsf{dontknowotherH} & := & \bigwedge_{a\neq b\in\{1,2,3\}} \bigwedge_{c\in\{r,w,b\}} (\neg c_a \rightarrow \neg K_a c_b) \end{array}$$

In dontknownotH, we may read $\neg K_a \neg$ as the dual epistemic modal operator meaning 'player a can imagine that' (M_a). All these properties hold in $Hexa$, and by corollary 4 therefore follow from $C_{123}\mathsf{Hexa}$ as well. Something more than that can be observed, however. We give an example: although seeH and dontseeH are not logically equivalent, we can replace seeH by dontseeH in theory Hexa and the result will *still* describe $Hexa$. This is captured by the following notion:

Definition 3 (Stronger) *Let Σ, φ, ψ be $\mathcal{L}_A^P$ formulas. Then 'φ* is *stronger than ψ* given *Σ', notation $\varphi \geq_\Sigma \psi$, if $C_A(\Sigma - \psi + \varphi) \models \psi$.*[3]

If either $\varphi \in \Sigma$ or $\psi \in \Sigma$, we say that '*φ is stronger than ψ* **in** Σ'. If $\varphi \geq_\Sigma \psi$ and $\psi \geq_\Sigma \varphi$ we write $\varphi =_\Sigma \psi$, and we say that '*φ and ψ are equally strong given/in* Σ' (or '*φ is just as strong as ψ*'). One can show that: $\mathsf{dealsH} =_{\mathsf{Hexa}} \mathsf{atmostH} \wedge \mathsf{atleastH}$, that $\mathsf{seeH} =_{\mathsf{Hexa}} \mathsf{dontseeH}$, that $\mathsf{dontknowthatH} =_{\mathsf{Hexa}} \mathsf{dontknowotherH}$, and that $\mathsf{dontknowthatH} =_{\mathsf{Hexa}} \mathsf{dontknownotH}$. Although the proofs need some combinatorial juggling, they are rather basic and have been omitted.

[3] $\Sigma - \psi = \Sigma \setminus \{\psi\}$ if $\psi \in \Sigma$ and $\Sigma - \psi = \Sigma$ otherwise. $\Sigma + \varphi = \varphi + \Sigma = \Sigma \cup \{\varphi\}$; whether φ is already in Σ or not makes no difference.

4 Description of initial game states

In the previous section we have described *Hexa*. In this section we generalize our results for any number of players and cards. We start by introducing terminology on deals of cards.

We consider deals $d \in A^{\mathbf{C}}$ of a set $\mathbf{C}$ of cards (in bold, to distinguish it from the common knowledge operator) over a finite set A of players. The cards held by player a can be represented by $d^{-1}(a) = \{c \in \mathbf{C} \mid d(c) = a\}$. Consider the game state where the cards have been dealt and every player has (only) looked at his own cards. In this *initial game state* two deals can not be distinguished from each other by a player (i.e., are the same for a player) if he holds the same cards in both *and* if all other players hold the same number of cards in both (imagine counting the backfaces of other players' cards). This induces an equivalence on card deals: *deal d is the same for player a as deal e* iff [$d^{-1}(a) = e^{-1}(a)$ and $\forall b \in A : |d^{-1}(b)| = |e^{-1}(b)|$]. We introduce some useful abbreviations. Let $a, b \in A$, $d \in A^{\mathbf{C}}$. Then $\sharp a := |d^{-1}(a)|$ is the number of cards held by player a. Similarly, write $\sharp\neg a$ for $|\mathbf{C}| - |d^{-1}(a)|$, the number of cards *not* held by a, and write $\sharp\neg ab$ for $|\mathbf{C}| - |d^{-1}(a)| - |d^{-1}(b)|$, the number of cards not held by a or b. The *size* of deal d, notation $\sharp d$, lists for each player the number of cards he holds.

In the following definition we represent the initial game state by an $S5$ model. The worlds in our model are deals. The atoms P of our language now take the form c_a, for $a \in A$ and $c \in \mathbf{C}$.

Definition 4 (Initial game state) *The initial game state is modelled by the pointed $S5_A$ model $(I_d, d) = (\langle D_{\sharp d}, \{\sim_a\}_{a\in A}, V\rangle, d)$, where $D_{\sharp d} = \{e \in A^{\mathbf{C}} \mid \forall a \in A : |e^{-1}(a)| = |d^{-1}(a)|\}$, and for all $a \in A$, $c \in \mathbf{C}$, $d_1, d_2 \in D_{\sharp d}$: $d_1 \sim_a d_2 \Leftrightarrow d_1^{-1}(a) = d_2^{-1}(a)$, and $V_{d_1}(c_a) = 1 \Leftrightarrow d_1(c) = a$.*

Definition 5 (Description of a deal of cards) *Let $d \in A^{\mathbf{C}}$. The* description of deal *d is the atomic description δ_d (see section 2) of the world d in I_d: $\delta_d := \bigwedge_{a\in A} \bigwedge_{c\in\mathbf{C}} c_a^d$. The* description of the cards of player *a is the part of that description about a: $\delta_d^a := \bigwedge_{c\in\mathbf{C}} c_a^d$.*

By definition we have that $\delta_d \leftrightarrow \bigwedge_{a\in A} \delta_d^a$. Also, for arbitrary $d', d'' \in I_d$ and $a \in A$: $\delta_d^a \leftrightarrow \bigvee_{d' \sim_a d} \delta_{d'}$, and $d \sim_a d' \Rightarrow \delta_d^a \leftrightarrow \delta_{d'}^a$. For more details on card deals and knowledge games, see (van Ditmarsch, 2001). We now present the theory Kgames:

Definition 6 (Kgames) *The theory* Kgames *for deal* $d \in A^{\mathbf{C}}$ *consists of the three constituents:*

$$\begin{array}{rcl} \mathsf{deals} & := & \bigvee_{d' \in D_{\sharp d}} \delta_{d'} \\ \mathsf{seedeal} & := & \bigwedge_{a \in A} \bigwedge_{d' \in D_{\sharp d}} (\delta^a_{d'} \rightarrow K_a \delta^a_{d'}) \\ \mathsf{dontknow} & := & \bigwedge_{a \in A} \bigwedge_{d' \in D_{\sharp d}} (\delta^a_{d'} \leftrightarrow \neg K_a \neg \delta_{d'}) \end{array}$$

Deals expresses that every world is atomically characterized by a deal of size $\sharp d$. Seedeal expresses that you know your own cards. Dontknow expresses that you can imagine any deal that is consistent with (in the sense of 'extends') your own cards. In a way, deals expresses the *factual knowledge* of all players, seedeal expresses the *private knowledge* of a player, and dontknow expresses his *private ignorance*. Again, all this is *publicly* known. In the coming subsections we discuss various other characterizations of the players' knowledge in I_d, and in what respect they are generalizations of constituents of Hexa. First, we show that Kgames describes I_d.

Theorem 5 *Theory* Kgames *for deal* $d \in A^{\mathbf{C}}$ *describes* I_d.

We have to show that I_d is a model of Kgames, which is obvious, and that *only* I_d models Kgames:

Lemma 6 *Suppose* $M \models$ Kgames. *Then* $M \underline{\leftrightarrow} I_d$.

Proof Let $M = \langle W^M, \{\sim^M_a\}_{a \in A}, V^M \rangle$ and $M \models$ Kgames; $I_d = \langle D_{\sharp d}, \{\sim_a\}_{a \in A}, V \rangle$. Observe that, because $M \models$ deals, each world $w \in M$ has a valuation $V_w = V_{d'}$ for some $d' \in D_{\sharp d}$. Define relation $\mathfrak{R} \subseteq (W^M \times D_{\sharp d})$ as follows: $\forall w \in W^M : \forall d' \in D_{\sharp d} : \mathfrak{R}(w, d') \Leftrightarrow V_w = V_{d'}$. We prove that $\mathfrak{R}$ is a bisimulation between M and I_d.

Atoms: By definition of $\mathfrak{R}$.

Forth: Let $w, w' \in M$, let $e \in D_{\sharp d}$. Suppose that $\mathfrak{R}(w, e)$ and that, for an arbitrary $a \in A$: $w \sim_a w'$. We find an $\mathfrak{R}$-image in $D_{\sharp d}$ of w' as follows:

Observe that $I_d, e \models \delta_e$. As $V_w = V_e$, also $M, w \models \delta_e$. Therefore $M, w \models \neg K_a \neg \delta_e$. From that and $M, w \models$ dontknow follows $M, w \models \delta^a_e$, i.e.: $M, w' \models \bigvee_{e' \sim_a e} \delta_{e'}$. Therefore there is an $e' \sim_a e$ such that $M, w' \models \delta_{e'}$. Choose one, then *that* e' is the required $\mathfrak{R}$-image of w': note that $e \sim_a e'$, and that $V^M_{w'} = V_{e'}$, because also, obviously, $I_d, e' \models \delta_{e'}$.

Back: Let $e, e' \in I_d$, let $w \in M$. Suppose that $\mathfrak{R}(w, e)$ and that, for an arbitrary $a \in A$, $e \sim_a e'$. We find an $\mathfrak{R}$-original of e', in M, as follows:

$M, w \models \delta_e$ implies, using seedeal, that $M, w \models K_a \delta_e^a$. Let $w'' \sim_a w$ be arbitrary. Then $M, w'' \models \delta_e^a$. From that and $e \sim_a e'$ follows $M, w'' \models \delta_{e'}^a$. Because w'' was arbitrary, $M, w \models K_a \delta_{e'}^a$. Therefore $M, w \models \delta_{e'}^a$. Using dontknow, we get $M, w \models \neg K_a \neg \delta_{e'}$. Therefore, there must be a $w' \sim_a w$ such that $M, w' \models \delta_{e'}$. Choose one, then *that* w' is the $\mathfrak{R}$-original of e', as $M, w' \models \delta_{e'}$ expresses that $V_{w'} = V_{e'}$. ■

Note that in the 'forth' part of the proof, we have only essentially used the $\neg K_a \neg \delta_e \rightarrow \delta_e^a$ part of dontknow, whereas in the 'back' part of the proof, we have essentially used seedeal and only used the $\delta_e^a \rightarrow \neg K_a \neg \delta_e$ part of dontknow.

Just as in the case of three players and three cards, that Kgames describes I_d means that we can neither *weaken* nor *strengthen* the theory, because I_d is its only model.

Corollary 7 *Theory* Kgames $+ \delta_e$ *describes* (I_d, e).

Corollary 8 *For all* $\varphi, \psi \in \mathcal{L}_A^P$ *(where* $P = \mathbf{C} \times A$*), and* $e \in I_d$*:* $I_d \models \varphi \Leftrightarrow I_d \models C_A\varphi \Leftrightarrow C_A$Kgames $\models \varphi \Leftrightarrow C_A$Kgames $\models C_A\varphi$*, and* $I_d, e \models \psi \Leftrightarrow$ Kgames $+ \delta_e \models \psi$.

In corollaries 7 and 8 it does not matter which point e we take in I_d, since $I_d = I_e$ as long as $\sharp d = \sharp e$. In particular, from corollary 7 it follows that the initial state (I_d, d) is described by Kgames $+ \delta_d$.

We continue with discussing various derived characteristics of the initial game state.

4.1 Derived characteristics: factual knowledge

Similarly to the case for $Hexa$, there are different ways to express how many cards players hold. As in section 3, we have that deals $=_{\mathsf{Kgames}}$ atmost $\wedge$ atleast.

Definition 7 (How many cards)

$$\begin{array}{lcl} \mathsf{atmost} & := & \bigwedge_{a \neq b \in A} \bigwedge_{c \in \mathbf{C}} \neg(c_a \wedge c_b) \\ \mathsf{atleast} & := & \bigwedge_{a \in A} \bigvee_{c^1 \neq \ldots \neq c^{\sharp a} \in \mathbf{C}} \bigwedge_{i=1}^{\sharp a} c_a^i \end{array}$$

4.2 Derived characteristics: private knowledge

We list six formulas describing what players 'see', that are all equally strong in Kgames.

Definition 8 (Seeing cards)

$$
\begin{array}{rcl}
\mathsf{see} & := & \bigwedge_{a\in A}\bigwedge_{c\in \mathbf{C}}(c_a \rightarrow K_a c_a) \\
\mathsf{dontsee} & := & \bigwedge_{a\in A}\bigwedge_{c\in \mathbf{C}}(\neg c_a \rightarrow K_a \neg c_a) \\
\mathsf{seeall} & := & \bigwedge_{a\in A}\bigwedge_{c^1\neq\ldots\neq c^{\sharp a}\in C}(\bigwedge_{i=1}^{\sharp a} c_a^i \rightarrow K_a \bigwedge_{i=1}^{\sharp a} c_a^i) \\
\mathsf{dontseeall} & := & \bigwedge_{a\in A}\bigwedge_{c^1\neq\ldots\neq c^{\sharp\neg a}\in C}(\bigwedge_{i=1}^{\sharp\neg a} \neg c_a^i \rightarrow K_a \bigwedge_{i=1}^{\sharp\neg a} \neg c_a^i) \\
\mathsf{seedeal} & := & \bigwedge_{a\in A}\bigwedge_{d'\in D_{\sharp d}}(\delta^a_{d'} \rightarrow K_a \delta^a_{d'}) \\
\mathsf{dontseedeal} & := & \bigwedge_{a\in A}\bigwedge_{d'\in D_{\sharp d}}(\neg\delta^a_{d'} \rightarrow K_a \neg\delta^a_{d'})
\end{array}
$$

See is the generalization of seeH, and dontsee of dontseeH: for every player and for every single card, if a player holds it, he knows that, and if he doesn't hold it, he knows that too. Instead, seeall and dontseeall express that, if a player holds his given number of cards, he knows them all, and, respectively, that if he does not hold any other card, he knows that too. Seedeal expresses that, if a player holds his given number of cards and does not hold any other card, he knows that. This is another way of saying that he knows his *local state* (see also section 6): for a given player a, to every atom c_a or $\neg c_a$ in δ^a_d corresponds the value of a 'local state variable for that player'. Dontseedeal expresses that if a player is *not* in a given local state, he knows that too. All six forms of 'seeing' are equally strong in Kgames. The proofs are simple and use deals. Although see is the most straightforward of all six, we prefer the more abstract formulation in constituent seedeal of Kgames.

4.3 Derived characteristics: private ignorance

The constituent dontknow of Kgames expresses private ignorance. It is conveniently formulated using the notations δ_d for the description of a deal and δ^a_d for the description of the cards of a. Directly reasoning from the properties of agents' knowledge, and starting from the characterization of private ignorance in $Hexa$ by dontknowthatH $=$ $\bigwedge_{a\neq b\in\{1,2,3\}}\bigwedge_{c\in\{r,w,b\}}\neg K_a c_b$, it is not so obvious how to get there. We illustrate this with an example:

Example 1 *Consider the initial game state for the game for three players 1, 2, 3 each holding two cards, with actual deal $kl|mn|op$. How ignorant is player 1 in this state of the game? Clearly, it is not strong enough that 1 does not know that 2 holds a specific combination of two cards: for all $c \neq c'$, $\neg K_1(c_2 \wedge c'_2)$. Player 1 also does not know that 2 holds any single card, which is stronger: for all c, $\neg K_1 c_2$. However, even that is not strong enough:*

Suppose 2 has told the others that he holds one of m and n. After that, it still holds that 1 doesn't know any of 2's cards: $\neg K_1 m_2$ and

$\neg K_1 n_2$. *However, 1 is less ignorant than before, because he* now *knows that 2 holds one of* two *cards:* $K_1(m_2 \vee n_2)$. *Initially, he didn't know that, indeed: for all* $c \neq c' \in \mathbf{C}$: $\neg K_1(c_2 \vee c'_2)$.

It appears that this is exactly the limit of his ignorance, because for some combinations of three *cards he* does *know that 2 holds one of them, e.g.* $K_1(m_2 \vee n_2 \vee o_2)$. *Suppose not, then it would be conceivable for player 1 that player 2 did not have any of the three cards* m, n, o. *Because player 1 holds two* other *cards himself,* k, l, *there would then be only one card left for player 2 to hold: card* p. *That player 2 holds only one card, contradicts* deals.

Definition 9 (Ignorance) *Let* $d \in A^{\mathbf{C}}$. *Then:*

$$\begin{array}{rcl} \mathsf{dontknowthat} & := & \bigwedge_{a\neq b\in A} \bigwedge_{c^1\neq\ldots\neq c^{\sharp\neg ab}\in C} \neg K_a \bigvee_{i=1}^{\sharp\neg ab} c_b^i \\ \mathsf{dontknownot} & := & \bigwedge_{a\neq b\in A} \bigwedge_{c^1\neq\ldots\neq c^{\sharp b}\in C} (\bigwedge_{i=1}^{\sharp b} \neg c_a^i \rightarrow \neg K_a \neg \bigwedge_{i=1}^{\sharp b} c_b^i) \\ \mathsf{dontknow} & := & \bigwedge_{a\in A} \bigwedge_{d'\in D_{\sharp d}} (\delta_{d'}^a \leftrightarrow \neg K_a \neg \delta_{d'}) \end{array}$$

The generalization of dontknowthatH reached in example 1 expresses characteristic dontknowthat in definition 9. Somewhat similarly, dontknownotH can be generalized to dontknownot. One can show that these two are equally strong in {deals, seedeal}. This is surprising, because dontknownot and dontknowthat appear to describe complementary kinds of ignorance: the first is on ignorance about cards other players have, the second is on ignorance about cards they *don't* have. However, both are still not strong enough to pin down I_d. We present another example to illustrate that:[4]

Example 2 *Consider the initial game state for the game for four players* $1, 2, 3, 4$ *each holding one card, with actual deal* $k|l|m|n$ *(klmn). Suppose an outsider tells player 1 that kmln is not the actual deal of cards, so that* $K_1 \neg \delta_{kmln}$. *Deal kmln is consistent with the local state, k, of player 1 in actual deal klmn, so that* $\neg K_1 \neg \delta_{kmln}$ *follows from* dontknow. *Therefore,* dontknow *does not hold in the current game state. However,* dontknowthat *holds: for any two cards, 1 can imagine that 2 does not have both. (Similarly,* dontknownot *holds.)*

We summarize the results: dontknowthat $=_{\{\mathsf{deals},\mathsf{seedeal}\}}$ dontknownot and dontknow $\geq_{\mathsf{Kgames}}$ dontknowthat (therefore also dontknow $\geq_{\mathsf{Kgames}}$ dontknownot). We show that dontknownot $\geq_{\{\mathsf{deals},\mathsf{seedeal}\}}$ dontknowthat. For the remaining proofs, see (van Ditmarsch, 2000):

[4] Suggested by Erik Krabbe.

Suppose not [dontknownot $\geq_{\{\text{deals},\text{see}\}}$ dontknowthat]. Then there are players a, b and cards $c^1, ..., c^{\sharp\neg ab}$ such that $K_a \bigvee_{i=1}^{\sharp\neg ab} c_b^i$. Regardless of whether a holds some of the cards $c^1, ..., c^{\sharp\neg ab}$, there must be at least $\sharp b$ other cards that a doesn't hold, suppose: $ca^1, ..., ca^{\sharp b}$. In other words, we have that: $\neg ca_a^1 \wedge ... \wedge \neg ca_a^{\sharp b}$. Applying dontknownot we get $\neg K_a \neg \bigwedge_{i=1}^{\sharp b} ca_b^i$. Formula $\neg K_a \neg \bigwedge_{i=1}^{\sharp b} ca_b^i$ means that a can imagine that b holds the $\sharp b$ cards $ca^1, ..., ca^{\sharp b}$. From $K_a \bigvee_{i=1}^{\sharp\neg ab} c_b^i$ it follows that a knows that b holds at least one more card, namely one of the (other!) cards $c^1, ..., c^m$. Therefore, a can imagine that b holds more than $\sharp b$ cards. From deals (that is known by a) follows that b holds exactly $\sharp b$ cards. Contradiction. ■

4.4 Derived characteristics: seedontknow

Using the validity of $K_a \delta_d^a \to \delta_d^a$ for $S5$ models, we can combine seedeal and dontknow in one formula seedontknow. It expresses that players can only imagine deals to be the case that correspond to what they know of their own cards.

Definition 10 (Combining private knowledge and ignorance)

$$\text{seedontknow} \quad := \quad \bigwedge_{a \in A} \bigwedge_{d' \in D_{\sharp d}} (K_a \delta_{d'}^a \leftrightarrow M_a \delta_{d'})$$

5 Description of the pre-initial state

We have described the initial state of a knowledge game, where the cards have been dealt and where players have looked at their cards. Now imagine that the cards have been dealt but that the players have not looked at their cards yet. Now players only know the *size* ($\sharp d$) of the deal: how many cards every player has. Therefore they consider every deal of that size a possibility. This is the *pre-initial state.*

Definition 11 (Pre-initial state) *The* pre-initial state *for a given deal* $d \in A^{\mathbf{C}}$ *is the state* $(preI_d, d)$, *where* $preI_d = \langle D_{\sharp d}, \{U_a\}_{a \in A}, V \rangle$ *with for each player* U_a *the universal relation* $D_{\sharp d} \times D_{\sharp d}$.

Definition 12 (Theory preKgames) *Let* $d \in A^{\mathbf{C}}$ *be a deal of cards. The theory* preKgames *(for deal* d*) consists of the constituents:*

$$\begin{aligned} \text{deals} &:= \bigvee_{d' \in D_{\sharp d}} \delta_{d'} \\ \text{dontknowany} &:= \bigwedge_{a \in A} \bigwedge_{d' \in D_{\sharp d}} \neg K_a \neg \delta_{d'} \end{aligned}$$

Theorem 9 *Theory* preKgames *for deal* $d \in A^{\mathbf{C}}$ *describes* $preI_d$.

We have to show that $preI_d$ is a model of preKgames, which, again, is obvious, and that *only* $preI_d$ models preKgames:

Lemma 10 *Suppose* $M \models$ preKgames. *Then* $M \underline{\leftrightarrow} preI_d$.

For a proof, see (van Ditmarsch, 2000). Just as in the corollaries 7 and 8, it follows that theory $\delta_e +$ preKgames describes $(preI_d, e)$, so that in particular $\delta_d +$ preKgames describes $(preI_d, d)$, and that $preI_d \models \varphi \Leftrightarrow preI_d \models C_A\varphi \Leftrightarrow C_A\text{preKgames} \models \varphi \Leftrightarrow C_A\text{preKgames} \models C_A\varphi$, and that preKgames $+ \delta_e \models \psi \Leftrightarrow preI_d, e \models \psi$.

6 Further observations

Fixed point computations We have described some $S5$ models (states) by way of proving that a suggested description (Hexa, Kgames, preKgames) indeed defines the bisimulation class of these models. These descriptions were 'merely' the outcome of a gradual process of generalizing properties of players' knowledge. They can also be directly computed, by a fixed-point construction that applies to finite modal states in general. See (van Benthem, 1998), relating to (Barwise and Moss, 1996). We have applied their construction to game states in (van Ditmarsch, 2000), and proven the results equivalent to the descriptions presented here.

Other game states We have described only the initial state of a knowledge game, and the pre-initial state. Even though we haven't given details on how to play knowledge games (for that see (van Ditmarsch, 2001)) it will be obvious that there are many other, possibly more complex, game states for such card games. Of course, future research should include the description of other game states. For many examples of game states, see (van Ditmarsch, 2000).

On the one hand we may expect *concise* descriptions of other game states to be more complex, because they will contain subgroup common knowledge operators. E.g. in *Hexa*, we have that $C_{12}r_1$ (1 and 2 commonly know that 1 holds the red card) is a postcondition of the action where 1 shows red to 2. On the other hand, such descriptions may be equivalent to publicly known theories with only occurrences of *individual* knowledge operators. E.g., in the given example it will suffice to state that K_2r_1 is one of three alternatives, that are commonly known.

Naturally, we would prefer not to have an ad hoc construction each time we are confronted with a new game state, whether this consists of constructing the unique model of a revised theory, or of devising the theory describing an intended model. We can improve on the situation in two ways, one semantically, the other syntactically.

In (van Ditmarsch, 2000), we present a dynamic epistemic language with dynamic modal operators $[\alpha]$ for actions α. In this action language we have described all knowledge game actions. An action is interpreted as a binary relation $[\![\alpha]\!]$ between states, in other words, as a semantic update. We have proven that action execution preserves bisimilarity of states. This provides us with an indirect – and much shorter – proof of lemma 10 that is needed for theorem 9: part of that indirect proof is to show that I_d results from executing an action α in $preI_d$, namely the action look_A corresponding to everybody looking at their cards (van Ditmarsch, 2000). This method could apply to relate other game states too.

A process of *syntactic relativization* of formulas (van Benthem, 2000b) may be preferable. There, a procedure is given for the special case of actions that are public announcements. If that procedure could be extended to game actions in general, we could 'directly' revise theories by a rewriting technique transforming constituents φ of the current description $\mathsf{Current}$ into constituents φ^α of the description Next of the next game state, such that expressions of type $\mathsf{Current} \rightarrow [\alpha]\mathsf{Next}$ are satisfied. For example, we would get $\mathsf{preKgames} \rightarrow [\mathsf{look}_A]\mathsf{Kgames}$. The factual knowledge of deals remains the same, i.e. $\mathsf{deals}^{\mathsf{look}_A} = \mathsf{deals}$, but $\mathsf{dontknowany} = \bigwedge_{a \in A} \bigwedge_{d' \in D_{\sharp d}} \neg K_a \neg \delta_{d'}$ will be relativized to $\mathsf{seedontknow} = \bigwedge_{a \in A} \bigwedge_{d' \in D_{\sharp d}} (K_a \delta^a_{d'} \leftrightarrow \neg K_a \neg \delta_{d'})$, i.e. $\mathsf{dontknowany}^{\mathsf{look}_A} = \mathsf{seedontknow}$. As the action look_A is not a public announcement the current procedure does not apply. We have not pursued this fascinating topic further.

Interpreted systems In section 4 we already paraphrased δ^a_d as the description of the *local state* of player a given deal d. In that sense, a deal of cards can be said to define a global state, with interesting dependencies between local states. Indeed, knowledge game states can be seen as interpreted systems in terms of (Fagin et al., 1995). This is summarily discussed in (van Ditmarsch, 2000), but also calls for further exploration. In particular, game state *frames* appear to have characteristics which correspond (in the technical sense) to multiagent axioms. Also this may help in describing other game states and models.

7 Conclusion

We have described two different states for card games. The model $Hexa$ for the initial game state of three players each holding a card is described by the theory Hexa. The model I_d of an initial game state for a deal d of $|\mathbf{C}|$ cards over $|A|$ players is described by the theory Kgames. Similarly, the model $preI_d$ where the players haven't looked at their cards yet is described by the theory preKgames. Descriptions of other game states may be fruitfully pursued from a viewpoint of knowledge relativization.

References

Barwise, J. and L. Moss. 1996. *Vicious Circles.* CSLI Publications, Stanford.

van Benthem, J. 1998. Dynamic odds and ends. Technical report, ILLC, University of Amsterdam. Report ML-1998-08.

van Benthem, J. 2000a. Hintikka self-applied. Manuscript.

van Benthem, J. 2000b. Information update as relativization. Manuscript.

Bonanno, G. 1993. The logical representation of extensive games. *International Journal of Game Theory*, 22(2):153–169.

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

van Ditmarsch, H. 2001. Knowledge games. *Bulletin of Economic Research.* To appear.

Ehrenfeucht, A. 1961. An application of games to the completeness problem for formalized theories. *Fundamenta Mathematicae*, 49:129–141.

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

Hintikka, J. and G. Sandu. 1989. Informational independence as a semantical phenomenon. In Fenstadt, J., I. Frolov, and R. Hilpinen, editors, *Logic, Methodology and Philosophy of Science VIII*, pages 571–589, Amsterdam. Elsevier.

Kaneko, M. and T. Nagashima. 1996. Game logic and its applications i. *Studia Logica*, 57:325–354.

Meyer, J.-J. and W. van der Hoek. 1995. *Epistemic Logic for AI and Computer Science.* Cambridge Tracts in Theoretical Computer Science 41. Cambridge University Press, Cambridge MA.