# What will they say? – Public Announcement Games

Thomas Ågotnes[*] Hans van Ditmarsch[†]

**Abstract**

Dynamic epistemic logics describe the epistemic consequences of actions. Public announcement logic, in particular, describe the consequences of public announcements. As such, these logics are *descriptive* – they describe what agents *can* do. In this paper we discuss what rational agents *will* or *should* do. We consider situations where each agent has a goal, a typically epistemic formula he or she would like to become true, and where the available actions are public announcements. What will each agent announce, assuming common knowledge of the situation? The truth value of the goal formula typically depends on the announcements made by several agents, hence we have a game theoretic scenario. We discuss possible solutions of such *public announcement games*.

## 1 Introduction

Dynamic epistemic logics describe the epistemic consequences of actions. Public announcement logic, in particular, describe the consequences of public announcements. As such, these logics are *descriptive* – they describe what agents *can* do, what pre- and post- conditions are, and so on. However, there is little *predictive* work in this area, describing what rational agents *will* do. In this paper we consider situations where each agent has a goal, a typically epistemic formula he or she would like to become true, and where the available actions are public announcements. What will each agent announce, assuming common knowledge of the situation? The truth value of the goal formula typically depends on the announcements made by several agents, hence we have a game theoretic scenario.

We make the following assumptions:

- agents have incomplete information about the world;
- agents have goals in the form of epistemic formulae, and agents' goals are common knowledge among all agents;

[*]University of Bergen, Norway, thomas.agotnes@infomedia.uib.no

[†]University of Sevilla, Spain, and University of Otago, New Zealand, hans@cs.otago.ac.nz

- each agent choose a (truthful) announcement (a formula she knows to be true);
- all agents make their announcements simultaneously; and
- all agents act rationally, i.e., they try to obtain their goals.

What can we say about how such agents will, or should, act?

In the next section we review the syntax and semantics of public announcement logic and some concepts from game theory. In Section 3 we introduce a formal model of *public announcement games*, and we discuss some possible solution concepts in Section 4 before we conclude in Section 5.

# 2 Background

## 2.1 Public Announcement Logic

The language $\mathcal{L}_{pal}$ of public announcement logic (PAL) [8] over a set of agents $N = \{1, \dots, n\}$ and a set of primitive propositions $\Theta$ is defined as follows, where $i$ is an agent and $p \in \Theta$:

$$\varphi ::= p \mid K_i\varphi \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid [!\varphi_1]\varphi_2$$

We write $\langle !\varphi_1\rangle\varphi_2$ resp. $\hat{K}_i\varphi$ for the duals $\neg[!\varphi_1]\neg\varphi_2$ and $\neg K_i\neg\varphi$.

A *Kripke structure* over $N$ and $\Theta$ is a tuple $M = (S, \sim_1, \dots, \sim_n, V)$ where $S$ is a set of states, $\sim_i \subseteq S \times S$ is an epistemic indistinguishability relation and is assumed to be an equivalence relation for each agent $i$, and $V : \Theta \to S$ assigns primitive propositions to the states in which they are true. A *pointed Kripke structure* is a pair $(M, s)$ where $s$ is a state in $M$. In this paper we will assume that Kripke structures are *finite* and *connected*.

The interpretation of formulae in a pointed Kripke structure is defined as follows (the other clauses are defined in usual truth-functional way).

$M, s \models K_i\varphi$ iff for every $t$ such that $s \sim_i t$, $M, t \models \varphi$

$M, s \models [!\varphi]\psi$ iff $M, s \models \varphi$ implies that $M|\varphi, s \models \psi$

where $M|\varphi = (S', \sim'_1, \dots, \sim'_n, V')$ such that $S' = \{s' \in S : M, s' \models \varphi\}$; $\sim'_i = \sim_i \cap (S' \times S')$; $V'(p) = V(p) \cap S'$.

The purely epistemic fragment of the language (i.e., formulae not containing public announcement operators $[!\varphi]$) is denoted $\mathcal{L}_{el}$. It was already shown in Plaza's original publication on that logic [8] that the language of PAL is no more expressive than the purely epistemic fragment.

## 2.2 Strategic Games

An *strategic game* is a tuple $G = \langle N, \{A_i : i \in N\}, \{u_i : i \in N\}\rangle$ where

- $N$ is the finite set of *players*
- for each $i \in N$, $A_i$ is the set of *strategies* (or *actions*) available to $i$. $A = \times_{j \in N} A_j$ is the set of *strategy profiles*.
- for each $i \in N$, $u_i : A \to \mathbb{R}$ is the *payoff function* for $i$, mapping each strategy profile to a number.

A strategy profile is a (pure strategy) *Nash equilibrium* if every strategy is the *best response* of that agent to the strategies of the other agents, i.e., if the agent can not do any better by choosing a different strategy given that the strategies of the other agents are fixed. A strategy for an agent is *weakly dominant* if it is as least as good for that again as any other strategy, no matter which strategies the other agents choose.

# 3 Public Announcement Games

Formally, a *public announcement game* models the agents' knowledge, and thereby available announcements, and goals:

**Definition 1 (Public Announcement Game)** *An ($n$-player) public announcement game (PAG) is a tuple*

$$AG = \langle M, \gamma_1, \dots, \gamma_n \rangle$$

*where $M$ is an epistemic structure, and $\gamma_i \in \mathcal{L}_{el}$ is* the goal formula *for agent $i$. A* pointed PAG *is a tuple $(AG, s)$ where $AG$ is a PAG and $s$ a state in $AG$. A* strategy *for agent $i$ in a pointed PAG is a formula $\varphi_i$ such that $M, s \models K_i \varphi_i$.*

It is now very natural to associate a strategic game with any pointed PAG $(AG, s)$: strategies, or actions, correspond to the individual announcements the agents can choose between, and a goal is satisfied iff it is true after all agents simultaneously make their chosen announcement. Formally:

**Definition 2 (State Game)** *The state game $G(AG, s)$ associated with state $s$ of PAG $AG = \langle M, \gamma_i, \dots, \gamma_n \rangle$ is defined by $N = \{1, \dots, n\}$, $A_i = \{\varphi_i : M, s \models K_i \varphi_i\}$ and*

$$u_i(\langle \varphi_1, \dots, \varphi_n \rangle) = \begin{cases} 1 & M, s \models \langle !K_1\varphi_1 \wedge \cdots K_n\varphi_n \rangle \gamma_i \\ 0 & \textit{otherwise} \end{cases}$$

Like in Boolean games [2, 7], binary utilities are implicit in public announcement games; an agent's goal is either satisfied or not.

**Counting strategies in the state game** A point to note is that all PAGs over finite epistemic structures can be *seen* as having only finitely many strategies, since there can be only finitely many announcements with different epistemic content. One can even enumerate those strategies, and the strategies can be described in a way independent from the actual state $s$.

Given a model $M$ with players $N = \{1, ..., n\}$, suppose that player $i$ has $m$ equivalence classes. We can show that the number of strategies for $i$ is $2^{m-1}$. Given that $i$ has $m$ equivalence classes, and that there are $2^m$ different subsets of a set of $m$ elements, there are $2^m$ unions of $i$-equivalence classes. Observe that the denotation of any announcement made by player $i$ *must* be such a union, as player $i$ only announces what she *knows* to be true: her announcement has form $K_i\varphi$. How many of those contain a given actual state? This number now is independent from that state. This is because for each such union of $i$-equivalence classes, its complement on the domain is also a union of $i$-equivalence classes; and any state should therefore be either in one or the other. In other words, we are counting the different ways to partition the domain in two parts (which we call a *dichotomy*) such that the partition induced by player $i$ is a refinement of that. Therefore there are $\frac{2^m}{2} = 2^{m-1}$ disjoint pairs of unions of $i$-equivalence classes.

In terms of announcements made by player $i$, a strategy for $i$ as defined above should be seen as an announcement *whether* player $i$ knows something, instead of the standard announcement *that* player $i$ knows something. The something the knowledge is about is the formula $K_i\varphi$, in other words, it is a (strictly speaking) non-deterministic event $!K_iK_i\varphi \cup !K_i\neg K_i\varphi$. Given positive and negative introspection this can be equated with non-deterministic event $!K_i\varphi \cup !\neg K_i\varphi$. (Or, alternatively, any other two exclusive known formulas considered to be a convenient description for the given game. See the example below.) As for a given state $s$ only of these can succeed, the event is functional and total. Therefore this is a proper strategy in the game theoretical sense.

**Example 3** *Consider the following formal model of a situation: a two-player pointed PAG* $(\langle M, \gamma_{Ann}, \gamma_{Bill}\rangle, s)$*, where* $M$ *is the following structure*

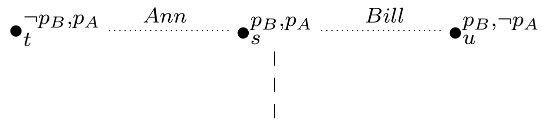
$$\bullet_t^{\neg p_B, p_A} \overset{Ann}{\cdots\cdots\cdots} \bullet_s^{p_B, p_A} \overset{Bill}{\cdots\cdots\cdots} \bullet_u^{p_B, \neg p_A}$$

*and*

$$\gamma_{Ann} = (K_B p_A \vee K_B \neg p_A) \rightarrow (K_A p_B \vee K_A \neg p_B)$$
$$\gamma_{Bill} = (K_A p_B \vee K_A \neg p_B) \rightarrow (K_B p_A \vee K_B \neg p_A)$$

*It is common knowledge that* $Ann$*'s goal is that* $Bill$ *does not get to know whether* $p_A$ *is true unless* $Ann$ *gets to know whether* $p_B$ *is true, and similarly the other way around. Actually,* $p_A$ *is true and* $Ann$ *knows this, and the same for* $p_B$ *and* $Bill$*.* $Ann$ *does not know whether* $Bill$ *already knows* $p_A$*, and similarly for* $Bill$*/*$Ann$*/*$p_B$*. Furthermore,* $Ann$ *does not know whether* $p_B$ *is true, but she knows that if* $p_B$ *is false then* $Bill$ *already knows that* $p_A$ *is true, and similarly for* $Bill$*.*

*In $s$ each agent can make two announcements with different information content, and the associated state game can thus be seen as a $2 \times 2$ matrix. We use the following picture to show that the game is associated with the point $s$:*



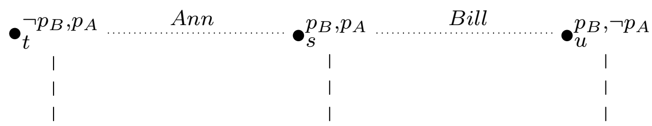

| | $\top$ | $\mathbf{p_B}$ |
|---|---|---|
| $\top$ | 11 | 10 |
| $\mathbf{p_A}$ | 01 | 11 |

*The figure above uses some notation we will use henceforth: $Ann$ is assumed to be the row player and $Bill$ the column player; payoff is written $xy$ where $x$ is $Ann$'s payoff and $y$ is $Bill$'s.*

*Notice that the game above has two Nash equilibria: either both agents announce their private information, or neither say anything informative. A winning strategy, for either agent, is to say nothing.*

*In terms of the enumeration of strategies that we introduced above, as both players have* two *equivalence classes, they both have $2^{2-1} = 2$ strategies. For example, Anne's partition is $\{t, s\}, \{u\}$. The four different unions of equivalence classes are $\emptyset, \{t, s\}, \{u\}, \{t, s, u\}$ and the two dichotomies are therefore $\{\emptyset, \{t, s, u\}\}$ and $\{\{t, s\}, \{u\}\}$. The first can be identied with the non-deterministic announcement $\neg K_A \top \cup K_A \top$, and therefore simply with announcement $\top$ (as this holds in any state), and the second with the non-deterministic announcement $!\neg K_A p_A \cup !K_A p_A$, which we can also see as the alternative between announcing $p_A$ or announcing $\neg p_A$ (in this example, if Ann does not know $p_A$ she knows $\neg p_A$). As we only describe the game for state $s$ where Ann indeed knows $p_A$, the second cannot actually occur. For Bill, we can similarly derive the alternatives $\top$ and $p_B$. This is wat we see in the figure.*

So a pointed PAG models the type of situations described in Section 1, and it might be tempting at first sight to view a pointed PAG similarly to a Boolean game, and use the game theoretic tool chest to define rational outcomes based on the state game. For example, in Example 3 we identified two Nash equilibria in the state game. However, observe that in state $s$ neither agent *knows* that the state actually is $s$ – and thus they do not necessarily know what the state game is! It is a fundamental assumption behind solution concepts such as the Nash equilibrium that the strategic game is common knowledge. Since the state game is not common knowledge among the two agents, the identification of equilibria of the state game can therefore not be a reliable method of identifying rational outcomes. Figure 1 illustrates the state games associated with also the two other states of Example 3. Clearly, if the actual state is $s$, the state game is not known by any of the players – in fact, they don't even know all the actions available to the other player. Indeed, while $(p_A, p_B)$ is a Nash equilibrium in the state game in $s$, it is not in the other




| | $\top$ | $\neg \mathbf{p_B}$ |
|---|---|---|
| $\top$ | 01 | 11 |
| $\mathbf{p_A}$ | 01 | 11 |

| | $\top$ | $\mathbf{p_B}$ |
|---|---|---|
| $\top$ | 11 | 10 |
| $\mathbf{p_A}$ | 01 | 11 |

| | $\top$ | $\mathbf{p_B}$ |
|---|---|---|
| $\top$ | 10 | 10 |
| $\neg \mathbf{p_A}$ | 11 | 11 |

Figure 1: PAG of Example 3 with state games.

state ($t$) which $Ann$ considers possible – she does not even know for certain that $p_B$ is a possible action for $Bill$.

Thus, the situations we are interested in can be modelled as a particular type of strategic games with imperfect information, where the strategies and information available in each state are closely interconnected (strategies *are* information) and where the same strategies are available in indiscernible states (but not necessarily in others). We now go on to discuss possible outcomes and solution concepts. In Section 5 we explain who standard models of strategic games with imperfect information, such as Bayesian games [3], can be applied to this setting.

# 4 Solution Concepts

Let us first consider weakly dominant (wd) strategies. It should be clear from the discussion above that there is a crucial distinction, in a pointed PAG, between:

- the *existence* of a strategy for an agent which is weakly dominant for that agent on the one hand, and
- the existence of a strategy for an agent which that agent *knows* is weakly dominant on the other,

because it might be the case that there is a strategy which is weakly dominant in one of the states she considers possible, but not in another. For example, in state $t$ of the model in Figure 1, $p_A$ is weakly dominant for $Ann$, but $Ann$ does not know this. Contrast this with the fact that $\top$ is also weakly dominant for $Ann$, but this $Ann$ knows.

Knowledge of weakly dominant strategies is a natural solution concept for PAGs. But another important distinction must be made.

**Definition 4** *Given a pointed PAG* $(\langle M, \gamma_i, \ldots, \gamma_n \rangle, s)$ *and an agent* $i$*, we say that* $i$ *has a* weakly dominant strategy de dicto *iff for any state* $t$ *such that* $s \sim_i t$*,* $i$ *has a weakly dominant strategy in the state game of* $t$*.*

**Definition 5** *Given a pointed PAG* $(\langle M, \gamma_i, \ldots, \gamma_n \rangle, s)$ *and an agent* $i$*, we say that* $i$ *has a* weakly dominant strategy de re *iff there is some strategy for* $i$ *which is weakly dominant in the state game of any state* $t$ *such that* $s \sim_i t$*.*

If an agent has a wd strategy *de dicto*, she knows *that* she has a wd strategy, i.e., she has a wd strategy in all states she considers possible, but she does not necessarily know *which* strategy is dominant; it is not necessarily the *same* strategy that is dominant in all the possible states. If she has a wd strategy *de re*, on the other hand, she knows *which* strategy is dominant; the same strategy is dominant in all the states she considers possible. Of course, having a wd strategy *de re* implies having one *de dicto*. The *de dicto*/*de re* distinction is well known in the knowledge and action literature [5, 4]. In state $s$ in the model in Figure 1, $Ann$ has a wd strategy *de re*, namely $\top$ ($Bill$ also has one – which one?). An example where an agent has a wd strategy *de dicto* but not *de re* will be shown later (Example 12).

What about the Nash equilibrium? Clearly, we can have similar situations: there might be a Nash equilibrium without the agents knowing it; the agents might know that there is a Nash equilibrium but not necessarily know what it is (there might be different equilibria in different accessible states). However, what "know" means here is not as clear as in the case of dominant strategies where knowledge of a single agent was needed. In the case of the Nash equilibrium there are *several* agents involved. Group notions of knowledge, such everybody-knows, distributed knowledge and common knowledge, have been studied in the context of the *de dicto*/*de re* distinction before [5]. For our purpose, we argue that the proper type of group knowledge for knowing a Nash equilibrium *de re* is *common knowledge*, since that is the assumption in game theory. Common knowledge of an equilibrium among all agents corresponds to a common equilibrium in all states of the model (since we assume connectedness). Thus, the existence of a Nash equilibrium *de re* is a *model* property, rather than a *pointed* model property, unlike existence of dominant strategies.

**Definition 6** *Given a PAG* $AG$*, we say that* there is a Nash equilibrium de re *if there exists a tuple of formulae, one for each agent, which constitutes a Nash equilibrium in the state game of every state in the PAG.*

For example, in the PAG in Figure 1, there is a Nash equilibrium *de re*, because the strategy profile $(\top, \top)$ is a Nash equilibrium in all the state games. An example where there are Nash equilibria in all the state games but no Nash equilibrium *de re* will be shown later (Example 12).

### 4.1 The Induced Game

Can a PAG be viewed as a (single) strategic game? We suggest the following definition.

**Definition 7** *Given a PAG* $AG = \langle M, \gamma_1, \ldots, \gamma_n \rangle$ *with* $M = (S, \sim_1, \ldots, \sim_n, V)$*, the* induced game $G_{AG}$ *is defined as follows:*

- $N = \{1, \ldots, n\}$
- *$A_i$ is the set of functions $a_i : S \to \mathcal{L}_{el}$ with the following properties:*
  - *Truthfulness: $M, s \models K_i a(s)$ for any $s$*
  - *Uniformity: $s \sim_i t \Rightarrow a_i(s) = a_i(t)$*

  *Thus, a strategy $a_i \in A_i$ gives a possible announcement for each state, but the same announcement for indiscernible states (note that the same announcements are always truthful in indiscernible states). Alternatively, $a_i$ can be seen as a function mapping* equivalence classes *to announcements.*
- *The payoffs are defined as follows. For any state $s$ in $AG$, let $G(AG, s) = (N, \{A_i^s : i \in N\}, \{u_i^s : i \in N\})$ be the state game associated with $s$ (Def. 2). Define, for any $(a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$:*

$$u_i(a_1, \ldots, a_n) = \frac{\sum_{s \in S} u_i^s(a_1(s), \ldots, a_n(s))}{|S|}$$

There are various important points to consider in the above definition.

**Strategies apply in all states** First, strategies are defined as plans for action in *any* possible state. This may look counter-intuitive if we want to find rational actions in some particular state of a PAG: agents know the available actions in that state (the same actions are available in all the states an agent considers possible). However, even though the current state is a member of the equivalence class one agent currently considers as possible states, she might consider many possibilities for what *another* agent's current equivalence class might be. Thus, she must take into account what the other agent is likely to do in all of these circumstances. Thus, a strategy must be description of behaviour for any contingency; even though each agent will only choose actions that actually are available in the current state.

**Common knowledge of the game** Second, payoff is computed by taking the average over *all states in the model*. It is clear that it does not suffice to look only in the current state, as each agent also might consider other states possible. But why not, then, compute an agent's payoff by taking the average over all the states that agent considers possible (the agent's equivalence class)? The reason is that the strategic game *must be common knowledge*, in order for solution concepts such that the Nash equilibrium to make sense. It might for example be that $Ann$ considers it possible that $Bill$ considers state $u$ possible, but that state $u$ is not in either $Ann$'s or $Bill$'s equivalence class for a current state $s$. If we take the average over only each agent's equivalence class for $s$, $u$ will not be taken into account. Averaging over all reachable states corresponds to averaging over all states commonly considered possible (all states accessible according to the accessibility relation for common knowledge). This is also the reason that the induced game is not induced from a *pointed* PAG: the induced game is the same at all points. This is as it should be, since the game should be common knowledge at any state. The computed payoffs can be seen as *expected* payoffs, not expected by a particular agent in the game, but expected payoffs as computed by a *common knower* – an agent whose knowledge is exactly the common knowledge among all agents in the game.

**Counting strategies in the induced game** We observed that in the state game for $(M, s)$ a player $i$ with $m$ equivalence classes has $2^{m-1}$ strategies, and that these strategies can be seen as taking the form of announcements *whether* some given formula is known by player $i$. In the induced game we can also count the number of strategies. The difference with the state game is that in a given state the player now can choose any of these $2^{m-1}$ state game strategies. Given uniformity, this means that player $i$ has that freedom for each of her $m$ equivalence classes. Of course, the action chosen in one equivalence class does not have to be the action chosen in another equivalence class: if Ann knows $p_A$, she might want to announce $\top$, but if Ann knows $\neg p_A$, she might not want to announce $\top$ but instead prefer to announce $\neg p_A$, etc. All these choices can be made completely independently. The total number of strategies for the induced game is therefore $2^{m-1} \times ... \times 2^{m-1}$ ($m$ times) which equals $(2^{m-1})^m = 2^{(m-1) \cdot m} = 2^{m^2 - m}$. This delivers a staggering number of strategy profiles $\times_{i \in \{1,...,n\}} 2^{m_i^2 - m_i}$ (where the number of equivalence classes for player $i$ is $m_i$). Given that the number of equivalence classes is in the order of the number of states $|M|$ of the model $M$, this gives us $O(2^{(|M| \cdot |M| \cdot n)})$ strategy profiles.

We will shortly explain the induced game further through several examples.

**Proposition 8** *If agent $i$ has a weakly dominant strategy* de re *in $(AG, s)$ for every state $s$ in a PAG AG, there is a weakly dominant strategy for $i$ in the induced game.*

**Proof** A weakly dominant strategy $a$ in the induced game is defined by taking $a(s)$ to be a wd strategy in the state game in $s$ and choosing the same strategy for all states int the same equivalence class (this is possible because the agent has a strategy *de re*). Wlog. assume that there are only two agents, and that $i = 1$. Suppose that $a$ is not weakly dominant. Then there is some other strategy $a'$ for 1, and some strategy $b$ for 2 such that

$$\frac{\sum_{s \in M} u_1^s(a'(s), b(s))}{|M|} > \frac{\sum_{s \in M} u_1^s(a(s), b(s))}{|M|}$$

Since payoffs are positive, this implies that $u_1^s(a'(s), b(s)) > u_1^s(a(s), b(s))$ for some $s$. But then $a(s)$ is not weakly dominant in the state game in $s$ after all, which is a contradiction. □

**Definition 9** *A* Nash Announcement Equilibrium *(NAE) of a PAG is a Nash equilibrium of the induced game.*

**Example 10** *Let us continue Example 3. We construct the induced game as follows (it is instructive to inspect the state games as illustrated in Fig. 1 on p. 6). $A_A$ (for $Ann$) contains the following four strategies:*

- $a_A^1$: $t, s \mapsto \top$; $u \mapsto \top$
- $a_A^2$: $t, s \mapsto \top$; $u \mapsto \neg p_A$
- $a_A^3$: $t, s \mapsto p_A$; $u \mapsto \top$
- $a_A^4$: $t, s \mapsto p_A$; $u \mapsto \neg p_A$

*$A_B$ (for $Bill$) is as follows:*

- $a_B^1$: $u, s \mapsto \top$; $t \mapsto \top$
- $a_B^2$: $u, s \mapsto \top$; $t \mapsto \neg p_B$
- $a_B^3$: $u, s \mapsto p_B$; $t \mapsto \top$
- $a_B^4$: $u, s \mapsto p_B$; $t \mapsto \neg p_B$

*In order to compute the payoffs, we need to check the payoffs in the state games for each state and combination of strategies. We have the following:*

| $\mathbf{a_A^x, a_B^y}$ | $\mathbf{t}$ | $\mathbf{s}$ | $\mathbf{u}$ |
|---|---|---|---|
| 1, 1 | 01 | 11 | 10 |
| 1, 2 | 11 | 11 | 10 |
| 1, 3 | 01 | 10 | 10 |
| 1, 4 | 11 | 10 | 10 |
| 2, 1 | 01 | 11 | 11 |
| 2, 2 | 11 | 11 | 10 |
| 2, 3 | 01 | 10 | 11 |
| 2, 4 | 11 | 10 | 11 |
| 3, 1 | 01 | 01 | 10 |
| 3, 2 | 11 | 01 | 10 |
| 3, 3 | 01 | 11 | 10 |
| 3, 4 | 11 | 11 | 10 |
| 4, 1 | 01 | 01 | 11 |
| 4, 2 | 11 | 01 | 11 |
| 4, 3 | 01 | 11 | 11 |
| 4, 4 | 11 | 11 | 11 |

*We get the following payoff matrix. We will henceforth write the payoffs* without *dividing by the number of states, for ease of presentation (the equilibria do of* *course not depend on this):*

| | $\mathbf{a_B^1}$ | $\mathbf{a_B^2}$ | $\mathbf{a_B^3}$ | $\mathbf{a_B^4}$ |
|---|---|---|---|---|
| $\mathbf{a_A^1}$ | $\underline{22}$ | $\underline{32}$ | 21 | 31 |
| $\mathbf{a_A^2}$ | $\underline{23}$ | 32 | 22 | 32 |
| $\mathbf{a_A^3}$ | 12 | 22 | $\underline{22}$ | $\underline{32}$ |
| $\mathbf{a_A^4}$ | 13 | 23 | $\underline{23}$ | $\underline{33}$ |

*The Nash equilibria are underlined.*

*Thus, the Nash announcement equilibria of this PAG are as follows, informally:*

**(1,1)** *Both agents say nothing (informative), no matter what*

**(1,2)** *$Ann$ says nothing, but $Bill$ says $\neg p_A$ if the state is $t$ (which $Bill$ can discern from any other state) and nothing otherwise. Let us consider this in the case that the current state is $s$. $Ann$ knows that the actual state is either $s$ or $t$, but not which. Thus, in the equilibrium she will play $\top$ under the assumption that $Bill$ will play $\top$ if the actual state is $s$ and $\neg p_A$ if the actual state is $t$ ($Bill$ can discern between these two possibilities). Actually, $Bill$ will play $\top$.*

**(2,1)** *Similarly, with $Ann$ and $Bill$ swapped*

**(3,3)** *$Ann$ says $p_A$ if she knows it, i.e., if the state is in $Ann$'s equivalence class $\{s, t\}$. Similarly for $Bill$.*

**(3,4)** *$Ann$ says $p_A$ if she knows it, and $Bill$ says $p_B$ if he knows it and $\neg p_B$ if he knows that*

**(4,3)** *Similarly, for $Ann$ and $Bill$ swapped*

**(4,4)** *Both agents say everything they know*

**Example 11 (Counting strategies)** *In Example 3, both Ann and Bill have two equivalence classes. An agent with $m$ equivalence classes has $2^{n^2-n}$ strategies in the induced game. Therefore, they both have $2^{2^2-2} = 4$ strategies. Indeed, this was observed in the previous example.*

*For another example, let player $A$ have three equivalence classes $x, y, z$, let $xy$ stand for the union of $x$ and $y$, etc.; then the four dichotomies are $xyz, x + yz, y + xz, z + xy$, and the total number of $A$ strategies for the induced game is $2^{3^2-3} = 2^6 = 64$. One of those strategies (entirely put with reference to structural properties of the model) is:*

*If $x$ then $xyz$, if $y$ then $y + xz$ (i.e.: $y$), and if $z$ then $xyz$.*

*Suppose, for a fairly typical arrangement, that $A$ knows $p$ in $x$, that $A$ knows $\neg p$ in $y$, and that $A$ is ignorant about $p$ in $z$. In other words, $x$ is the denotation of $K_A p$, $y$ is the denotation of $K_A \neg p$, and $z$ is the denotation of $\neg(K_A p \vee K_A \neg p)$ (which is equivalent to the known formula $K_A \neg(K_A p \vee K_A \neg p)$). Then one of those 64 strategies, employing logical notation instead, is:*

> *if $K_A p$ then $!T$, if $K_A \neg p$ then $(!K_A \neg p \cup !\neg K_A \neg p)$, and if $\neg(K_A p \vee K_A \neg p)$ then $!T$.*

*which in natural language becomes*

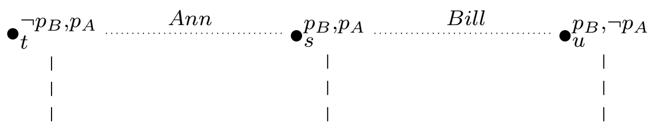
> *if I know $p$ then I make the trivial announcement, if I know $\neg p$ then I announce $\neg p$, and if I am ignorant about $p$ then I also make the trivial announcement.*

**Example 12** *Define a PAG AG as follows. Let the model be as in Example 3, but change the goals as follows:*

$$\gamma_{Ann} = (K_B(p_A \wedge p_B) \wedge \neg K_A p_B) \vee (K_B(\neg p_B \wedge p_A) \wedge \hat{K}_A \hat{K}_B \neg p_A) \vee (K_A(p_B \wedge \neg p_A) \wedge \hat{K}_B \hat{K}_A \neg p_B)$$

$$\gamma_{Bill} = (K_A(p_A \wedge p_B) \wedge \neg K_B p_A) \vee (K_B(\neg p_B \wedge p_A) \wedge \hat{K}_A \hat{K}_B \neg p_A) \vee (K_A(p_B \wedge \neg p_A) \wedge \hat{K}_B \hat{K}_A \neg p_B)$$

*Perhaps the reader finds these long formulae hard to read, but it suffices to trust that they give the following state games:*




| | $\top$ | $\neg \mathbf{p_B}$ |
|---|---|---|
| $\top$ | 11 | 00 |
| $\mathbf{p_A}$ | 00 | 00 |

| | $\top$ | $\mathbf{p_B}$ |
|---|---|---|
| $\top$ | 00 | 01 |
| $\mathbf{p_A}$ | 10 | 00 |

| | $\top$ | $\mathbf{p_B}$ |
|---|---|---|
| $\top$ | 11 | 00 |
| $\neg \mathbf{p_A}$ | 00 | 00 |

*The PAG has some properties not found in the PAG in Example 3 (Figure 1). First, $Ann$ has a weakly dominant strategy* de dicto*, but not* de re*, in the pointed PAG $(AG, s)$. The strategy $p_A$ is weakly dominant in $s$, but not in $t$. There is, however, another weakly dominant strategy in $t$, namely $\top$. Second, while every state game has a Nash equilibrium, there does not exist a Nash equilibrium* de re *in $AG$.*

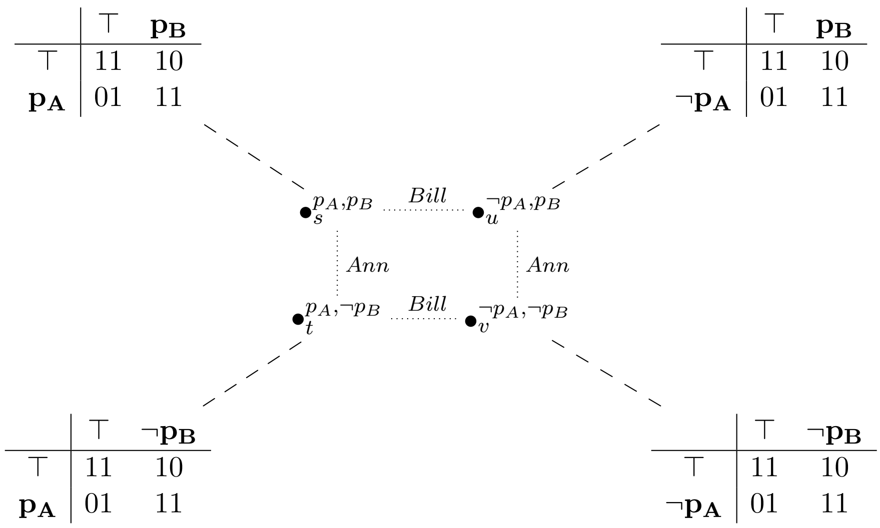
*We get the following induced game, where the strategies are as in Example 10:*

| | $\mathbf{a_B^1}$ | $\mathbf{a_B^2}$ | $\mathbf{a_B^3}$ | $\mathbf{a_B^4}$ |
|---|---|---|---|---|
| $\mathbf{a_A^1}$ | $\underline{22}$ | 11 | $\underline{12}$ | 01 |
| $\mathbf{a_A^2}$ | 11 | 00 | $\underline{12}$ | 01 |
| $\mathbf{a_A^3}$ | $\underline{21}$ | $\underline{21}$ | 00 | 00 |
| $\mathbf{a_A^4}$ | 10 | 10 | 00 | $\underline{00}$ |

*The Nash announcement equilibria are underlined. Let us consider the situation in state $s$. There are several different Nash announcement equilibria, including: all agents announce $\top$ in all states (including in $s$). Note that $(\top, \top)$ is* not *a Nash equilibrium in the state game in $s$. Another equilibrium is that both agents play "down" (i.e., not $\top$) in any state (also a Nash equilibrium in the state game in $s$). Note that $(a_A^4, a_B^4)$* is *a NAE while for example $(a_A^3, a_B^4)$ is* not. *If the current state is $s$ (or, from $Ann$'s perspective the current equivalence class is $\{s, t\}$ and* *from Bill's perspective $\{s, u\}$), $Ann$ will in fact do exactly the same if she uses strategy $a_A^3$ or $a_A^4$. However, since $Bill$ does not know whether $Ann$'s equivalence class is $\{s, t\}$ or $\{u\}$, he must also consider what $Ann$ does in $u$ – which is exactly what differentiates $a_A^3$ and $a_A^4$. Thus, the distinction between these two strategies is significant, even in a state ($s$) where they give the same action.*

**Example 13** *Let us consider a more regular and symmetric PAG than the ones discussed so far. The situation is similar to the one in Example 3, but now $Ann$ knows that $Bill$ does not know $p_A$, and similarly for $Bill$/$Ann$/$p_B$. The situation is modelled by the following goal formulae and Kripke structure. We have also shown the state games.*

$$\gamma_{Ann} = (K_B p_A \vee K_B \neg p_A) \rightarrow (K_A p_B \vee K_A \neg p_B)$$
$$\gamma_{Bill} = (K_A p_B \vee K_A \neg p_B) \rightarrow (K_B p_A \vee K_B \neg p_A)$$




*Again, the induced game has four distinct strategies for each agent:*

| $\mathbf{x}$ | $\mathbf{a_A^x}$ | | $\mathbf{a_B^x}$ | |
|---|---|---|---|---|
| 1 | $s, t \mapsto \top$; | $u, v \mapsto \top$ | $s, u \mapsto \top$; | $t, v \mapsto \top$ |
| 2 | $s, t \mapsto \top$; | $u, v \mapsto \neg p_A$ | $s, u \mapsto \top$; | $t, v \mapsto \neg p_B$ |
| 3 | $s, t \mapsto p_A$; | $u, v \mapsto \top$ | $s, u \mapsto p_B$; | $t, v \mapsto \top$ |
| 4 | $s, t \mapsto p_A$; | $u, v \mapsto \neg p_A$ | $s, u \mapsto p_B$; | $t, v \mapsto \neg p_B$ |

*The induced game (Nash equilibria underlined):*

| | $\mathbf{a_B^1}$ | $\mathbf{a_B^2}$ | $\mathbf{a_B^3}$ | $\mathbf{a_B^4}$ |
|---|---|---|---|---|
| $\mathbf{a_A^1}$ | $\underline{44}$ | 42 | 42 | 40 |
| $\mathbf{a_A^2}$ | 24 | 33 | 33 | 42 |
| $\mathbf{a_A^3}$ | 24 | 33 | 33 | 42 |
| $\mathbf{a_A^4}$ | 04 | 24 | 24 | $\underline{44}$ |

*The game has two Nash equilibria. The first is that both agents say nothing,* in all states. *The strategies in this equilibrium are both dominant strategies. The second equilibrium ($a^4_A, a^4_B$) is that both agents tell everything they know,* in all states.

In Example 13 the Nash announcement equilibria are all "composed" of Nash equilibria in the state game, in the following sense: for every NAE $(a, b)$ and every state $s$, $(a(s), b(s))$ is a Nash equilibrium in the state game in $s$ (albeit not *all* such compositions of Nash equilibria in the state games are NAE in the example). Indeed, this is also the case in Example 3. Is this a general property of PAGs? No, and a counter example is found in Example 12: $(a^1_A, a^1_B)$, because $(a^1_A(s), a^1_B(s))$ is not a Nash equilibrium in the state game in $s$.

We can establish a connection to having a Nash equilibrium *de re*, similarly to Proposition 8 for dominant strategies.

**Proposition 14** *If there exists Nash equilibrium* de re *in a PAG, then there exists a Nash announcement equilibrium.*

**Proof** Assume wlog. that there are only two agents. If there is a Nash equilibrium *de re*, then there is a strategy profile $(x, y)$ which is a Nash equilibrium in every state game. Let $(a, b)$ be a strategy profile for the induced game such that $a(s) = x$ and $b(s) = y$ for any $s$. Clearly, $a$ and $b$ are both uniform and truthful. Suppose that $(a, b)$ is not a Nash equilibrium in the induced game. Then there is a better response $a'$ for one of the agents, again wlog. assume for agent 1. In other words, there is a strategy $a'$ for agent 1 such that $u_1(a', b) > u_1(a, b)$. But this entails that $u^s_1(a'(s), b(s)) > u^s_1(a(s), b(s))$ for some state $s$, and thus that there is a strategy $z$ for agent 1 in the state game in $s$ such that $u^s_1(z, y) > u^s_1(x, y)$ – which contradicts the fact that $(x, y)$ is a Nash equilibrium in the state game in $s$. □

Proposition 14 does not hold in the other direction. A counter example is found in Example 12.

## 5 Discussion

**Boolean games** The intimate connection between knowledge and strategies in public announcement games distinguishes them from many other types of games. In *Boolean games* [2, 7], each agent has a goal formula like in PAGs, and each agent controls a set of primitive propositions which affects the truth value of the goal formulas. In contrast, in PAGs an agent "controls" common knowledge of any formula he or she knows. We have seen that we cannot simply view a pointed PAG as Boolean type game, because the agents do not necessarily have common knowledge about the game that is being played.

**Bayesian games** The most common model of strategic games with imperfect information is the *Bayesian game* [3]. Public announcement games can be seen as Bayesian games. We base our comparison on the excellent presentation of Bayesian games in [6]. This comparison is fairly straightforward. Given a parameter model $M$ for a public announcement game, the required finite set of states for the Bayesian game is the domain of that model. A *signal* observed by a player $i$ in the Bayesian game corresponds to the *equivalence class* of that player in the model $M$, and the signal function mapping a state to the signal for player $i$ in that state therefore maps each state/world $s$ in the domain of $M$ to the equivalence class $[s]_{\sim_i}$ for the player $i$ (where different players are assigned different classes).

Next the somewhat more tricky issue of the probability measure. In Bayesian games it is assumed that each player has a probability measure over the set of states. This can easily be adapted to the possible worlds framework: we assume that all states/worlds in $M$ get equal probability $\frac{1}{|M|}$. This is justified because the structure of the model $M$ is assumed to be commonly known to all players, and because no player has any a priori reason to prefer any state in $M$ over any other. Consider a typical example of an epistemic model $M$. Given a deal of cards over $n$ players where the players have been dealt their cards but where they have not picked up their cards yet (and where the cards are lying face down on the table in front of them), no player has any reason to prefer any particular deal over another; no player has any reason to assume that she is more likely to get the ace of spades than the three of hearts before picking up her cards (unless cheating or unfair play is the case). In other words, an a priori equal distribution of probability over all states in a Kripke model is a fairly standard assumption in the knowledge games that we investigate. Similarly, the a posteriori distribution of probabilities for a player $i$, after receiving her signal, is conditionalized over the states in her equivalence class: once you picked up your cards, you no longer consider it possible that you have other cards in your hands.

The actions (strategies) available to a player $i$ in the Bayesian game correspond to the non-deterministic announcements (player $i$ announces whether she knows some epistemic formula) that we defined as strategies for the state public announcement game. For the purpose of computing equilibria of the induced public announcement game we defined strategies somewhat differently, namely as *conditionalized actions* such as 'If $i$ knows $p$, then $i$ announces $p$, and if $i$ knows $\neg p$, then $i$ makes the trivial announcement'; in other words, given the requirement for uniform strategies, in the induced PAG a pure strategy for player $i$ is a total function from signals for $i$ ($i$-equivalence classes) to announcements by $i$ (actions). In Bayesian games the relation between actions and players is modelled differently: to compute an equilibrium for a Bayesian game we define the game where each combination $(i, t_i)$ of a player $i$ and a signal (equivalence class in our case, therefore:) $t_i = [s]_{\sim_i}$ defines a virtual player. The equilibria are then computed for the game with those virtual players. For example, instead of the single strategy above we would get two strategies, for different players, namely (replacing the equivalence class by the formula denoting it) 'Player $(i, K_i p)$ plays (announces) $p$', 'Player $(i, K_i \neg p)$ plays (announces) the trivial announcement'. For pure strategy equilibria this does not make a difference between Bayesian games and induced public announcement games in our modelling.

For mixed equilibria this seems less clear, but we conjecture that the match with Bayesian games is perfect, and that also the mixed equilibria should be identical for both modellings. Here, we can illustrate the problem by our example. Consider a mixed strategy equilibrium for an induced PAG where for player $i$ the strategy 'If $i$ knows $p$, then $i$ announces $p$, and if $i$ knows $\neg p$, then $i$ makes the trivial announcement' has weight $0.4$. It seems (naively – we are thinking ahead to our solution, to follow) that the corresponding mixed strategy for the Bayesian modelling should have for player $(i, K_i p)$ the strategy '(announce) $p$' with weight $0.4$ and for (different) player $(i, K_i \neg p)$ strategy '(announce) the trivial announcement' with weight $0.4$ as well. But in the Bayesian modelling pure strategies for different players can of course also have *different* weights, e.g., for player $(i, K_i p)$ the strategy '(announce) $p$' with weight $0.6$ and for player $(i, K_i \neg p)$ strategy '(announce) the trivial announcement' weight $0.3$. What would correspond to that in the induced public announcement game? We conjecture that a one-to-one correspondence can be defined: in order to determine the latter we need to take all induced PAG strategies into account that have a conditional part 'If $i$ knows $p$, then $i$ announces $p$' and also all induced PAG strategies that have another conditional part 'if $i$ knows $\neg p$, then $i$ makes the trivial announcement'. And there are more of those than the single pure strategy 'If $i$ knows $p$, then $i$ announces $p$, and if $i$ knows $\neg p$, then $i$ makes the trivial announcement', for example, another one is 'If $i$ knows $p$, then $i$ makes the trivial announcement, and if $i$ knows $\neg p$, then $i$ makes the trivial announcement'. We intend to resolve this issue in future research.

**Question-answer games** Consider a different game. Instead of players choosing what to announce, players choose what questions to pose to another player, where the other player is obliged to answer the question truthfully. Of course there are just as many questions to be posed to player $j$ as (truly) different announcements for player $j$ to make. And one variable in such a game could be that a player $i$ whose turn it is may *choose* another player $j$ to ask a question to. We can already observe that, given such a choice, the total number of strategies of the induced question-answer game is also countable given some model $M$ involving at least players $i$ and $j$ as initial parameters. Instead of $2^{m^2-m}$ pure strategies for player $i$ with $m$ equivalence classes, we now have $2^{m_j m_i - m_i}$ pure strategies for player $i$ asking player $j$ a question, where player $i$ has $m_i$ equivalence classes and player $j$ has $m_j$ equivalence classes. This we can see as follows. There are $2^{m_j-1}$ different dichotomies for player $j$ (i.e. coursenings of player $j$'s partition), and for each of $m_i$ different equivalence classes for the requesting player $i$, she may choose one of those questions, therefore the total number of pure strategies is $(2^{m_j-1})^{m_i} = 2^{m_j m_i - m_i}$.

This sort of question-answer game is defined as a knowledge game in van Ditmarsch PhD thesis [9], for the more general case where the question is public but the answer may be semi-public: the other players know what the question is, but may only partially observe the answer (e.g., the question may be to show a card *only* to the requesting player but where there is common knowledge of some card being shown; the alternatives are the different cards to be shown, from our current perspective a subset of all different questions to be asked namely only those with singleton equivalence classes). These matters is also addressed in [10, 11].

Stefan Minica, ILLC, University of Amsterdam, is currently investigating such question-answer games in collaboration with us.

**Further generalizations** Is our definition of the Nash announcement equilibrium the right one? We have argued that it is reasonable and has desirable properties, e.g., it is an equilibrium of a game that is common knowledge among all agents, and the payoffs are expected payoffs computed by a "common knower". Further studies of its properties are needed, and this is work in progress. In future work we will also study mixed strategies, we will look at more fine grained goal models which do not necessarily give binary payoffs, for example lists of prioritised goals [1], and we will model situations with *sequential* announcements by using extensive form games.

# References


[1] Thomas Ågotnes, Michael Wooldridge, and Wiebe van der Hoek. Normative system games. In M. Huhns and O. Shehory, editors, *Proceedings of the Sixth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007)*, pages 876–883. IFAMAAS, May 2007.

[2] P. Harrenstein. *Logic in Conflict*. PhD thesis, Utrecht University, 2004.

[3] J. C. Harsanyi. Games with Incomplete Information Played by 'Bayesian' Players, Parts I, II, and III. *Management Science*, 14:159–182, 320–334, 486–502, 1967–1968.

[4] Wojciech Jamroga and Thomas Ågotnes. Constructive knowledge: what agents can achieve under imperfect information. *Journal of Applied Non-Classical Logics*, 17(4):423–475, 2007.

[5] Wojciech Jamroga and Wiebe van der Hoek. Agents that know how to play. *Fundamenta Informaticae*, 63:185–219, 2004.

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

[7] J.-J.Ch. Meyer P. Harrenstein, W. van der Hoek and C. Witteveen. Boolean games. In J. van Benthem, editor, *Proceeding of the Eighth Conference on Theoretical Aspects of Rationality and Knowledge (TARK VIII)*, pages 287–298, Siena, Italy, 2001.

[8] J.A. Plaza. Logics of public communications. In M.L. Emrich, M.S. Pfeifer, M. Hadzikadic, and Z.W. Ras, editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems: Poster Session Program*, pages 201–216. Oak Ridge National Laboratory, 1989. ORNL/DSRD-24.

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

[10] H.P. van Ditmarsch. The description of game actions in cluedo. In L.A. Petrosian and V.V. Mazalov, editors, *Game Theory and Applications*, volume 8, pages 1–28, Commack, NY, USA, 2002. Nova Science Publishers.

[11] H.P. van Ditmarsch. Some game theory of pit. In C. Zhang, H.W. Guesgen, and W.K. Yeap, editors, *Proceedings of PRICAI 2004 (Eighth Pacific Rim International Conference on Artificial Intelligence)*, pages 946–947. Springer, 2004. LNAI 3157.