Awareness in repeated games

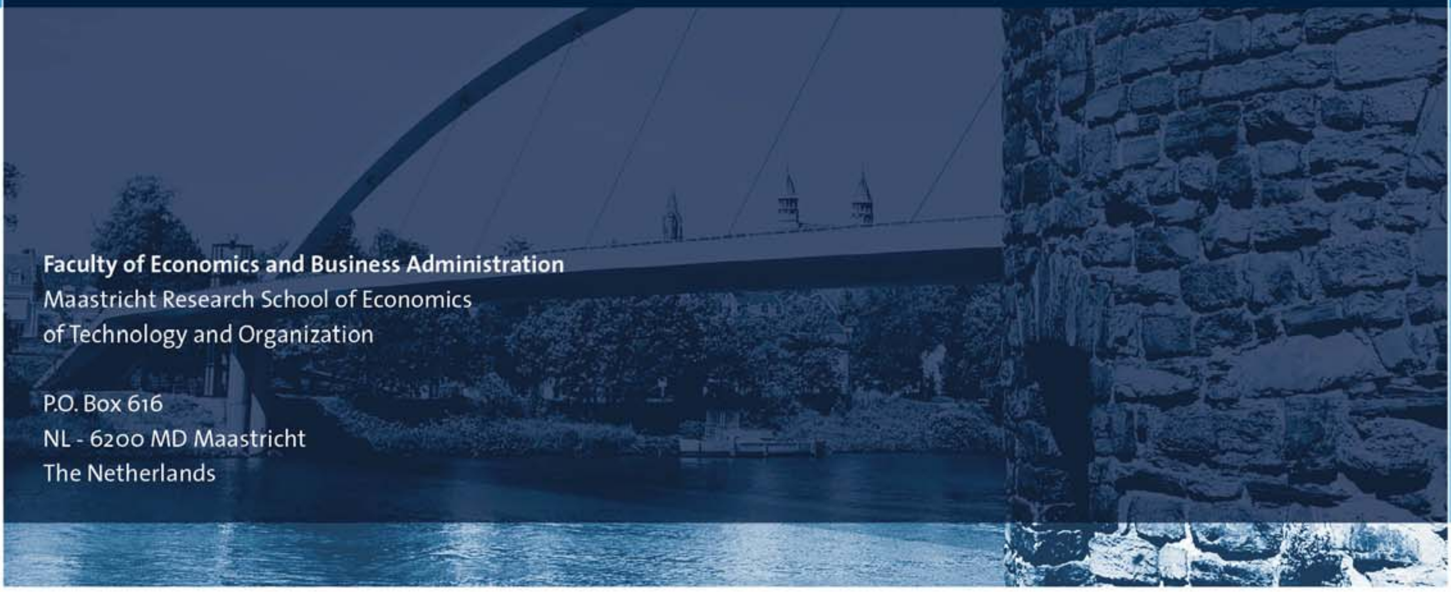
Maastricht University

Friederike Mengel, Elias Tsakas,
Alexander Vostroknutov

Awareness in Repeated Games

RM/09/010

METEOR



Faculty of Economics and Business Administration
Maastricht Research School of Economics
of Technology and Organization

P.O. Box 616
NL - 6200 MD Maastricht
The Netherlands

Awareness in repeated games*

FRIEDERIKE MENGEL,[†] ELIAS TSAKAS,[‡] ALEXANDER VOSTROKNUTOV[§]

Department of Economics, Maastricht University

February, 2009

Abstract

In this paper we provide a framework to reason about limited awareness of the action space in finitely repeated games. Our framework is rich enough to capture the full strategic aspect of limited awareness in a dynamic setting, taking into account the possibility that agents might want to reveal or conceal actions to their opponent or that they might become "aware of unawareness" upon observing non rationalizable behavior. We show that one can think of these situations as a game with incomplete information, which is fundamentally different, though, from the standard treatment of repeated games with incomplete information. We establish conditions on the "level of mutual awareness" of the action space needed to recover Nash and subgame perfect Nash equilibria from the standard theory with common knowledge. We also show that the set of sustainable payoffs in games with folk theorems does not relate in a monotone way to the "level of mutual awareness".

KEYWORDS: Repeated Games, Unawareness, Discovery, Rationalizability.

JEL CLASSIFICATION: C70, C72, D80, D82

*We are indebted to Jayant Ganguli, Ronald Peeters and Andrés Perea for helpful comments and fruitful discussions.

[†]Department of Economics (AE1), Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands; Tel:+31-43-38 83891; Fax:+31-43-38 84878; E-mail: f.mengel@algec.unimaas.nl

[‡]Department of Economics (AE1), Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands; Tel:+31-43-38 83649; Fax:+31-43-38 84878; E-mail: e.tsakas@algec.unimaas.nl

[§]Department of Economics (AE1), Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands; Tel:+31-43-38 84606; Fax:+31-43-38 84878; E-mail: a.vostroknutov@algec.unimaas.nl

1. Introduction

While standard game theory implicitly assumes that players can reason about all aspects of a game, this assumption is certainly too strong in some real life strategic interactions. As an example consider a firm procuring maintenance services from the company they bought some complicated machines from. The company tells them that there are essentially two maintenance options A and B and they agree on a contract specifying the price for each of the two options. After starting the maintenance work, though, the company makes an unexpected announcement to the firm telling them that the situation is more complex and that now the best option would be option C, a contingency the firm was initially unaware of.[1] Note, though, that the second time the firm procures maintenance services it will not be unaware of possibility C. Other examples include financial markets where some investors might not be aware of some (complex) investment options, tax avoidance possibilities etc...

Much of the strategic interest in these examples resides in the fact that awareness might change over time. Just as the firm becomes aware of option C after observing it once, investors in financial markets will learn about an action they were previously unaware of once one of their competitors has successfully applied it. Of course this raises the question whether and when an investor wants to reveal such options to other investors. The dynamic model also raises another issue. Think of an investor i that observes a competitor j choosing an action which he is aware of, but which seems completely irrational to him. Such an investor might become "suspicious" or "aware of unawareness". In this case investor i might want to make an effort to "discover" new options, since he suspects to be unaware of something. But then again if investor j knows that i might become suspicious shouldn't it be optimal for him to refrain from choosing such an option ?

These examples suggest that by neglecting the dynamic aspect of limited awareness one misses out on a number of interesting strategic phenomena. A complete theory of belief and strategic choices should thus allow for the discovery of new, previously unconsidered propositions, actions or states of nature and provide a framework to analyze the whole range of strategic implications that follow. Existing literature, though, has largely neglected the interesting strategic issues that arise from the essentially dynamic structure of unawareness. Few exceptions are Heifetz, Meier and Schipper (2009) as well as Halpern and Rego (2006) or Grant and Quiggin (2007), all of which we will discuss below.

In this paper we take an explicitly dynamic approach that is able to capture all the phenomena discussed above. We model agents interacting in finitely repeated two-player games allowing for

[1] Lee (2008) among others discusses contracting with unforeseen contingencies.

the possibility that one (or both) agents might not be aware of the entire action set.[2] Naturally in such a dynamic setting the possibility arises that agents might become aware of (or "discover") actions they were previously unaware of. An obvious instance where this occurs is if they observe the opponent choose an action they were previously unaware of. Such considerations entail the question of whether agents can be "aware of unawareness" and how and whether this can be reconciled with the rationality of agents. Unlike Grant and Quiggin (2007) we try to model "discovery" of new actions by staying as close as possible to the principle of rationality. We assume that agents are "confident", meaning that - as long as all observed behavior is rationalizable given their current awareness - they are not aware of the fact that they might be unaware of something. If they do observe something, though, that is not rationalizable for them we say they become "aware of unawareness" and only in this case they might discover actions they were not previously aware of.

Of course agents can also reason about their opponent's awareness and hence form beliefs about it. In fact the entire hierarchy of beliefs has to be considered. We use a set theoretic structure to model unawareness and show that one can think of this problem as a (repeated) game of incomplete information where each agent's type describes his current awareness together with a hierarchy of beliefs about his opponent's awareness. Such a type space, though, cannot capture the essentially dynamic nature of the problem. We thus introduce a type space for the dynamic setting where each type includes a player i's awareness, beliefs about the other player j's awareness, beliefs about the beliefs of j about the awareness of i etc.. *at all contingencies* of the game. This setting is no longer equivalent to a standard incomplete information setting (for repeated games), since players might not initially be aware of all their information sets. Such unforeseen contingencies might lead to changes in the agent's awareness and induce e.g. the possibility that an agent's current belief about the opponent's future type is correlated with his action choice.[3] We show that this difference is far from innocuous and has implications for equilibrium play that we will discuss below. We impose some restrictions on types, such as coherency (a player cannot believe that his opponent is aware of more things than himself), perfect recall and few others and show that the resulting type space is well behaved. We define what rationalizable behavior is in this context, define Nash (and subgame perfect Nash) equilibrium and prove their existence.

[2] The question might come to mind why we consider limited awareness of the action space and not of some repeated game strategies. Given full awareness of the action space the latter would essentially mean a departure from rationality. Since we want to distinguish between awareness and rationality we assume that our players are fully rational (in the standard game theoretic sense) given their awareness.

[3] Other differences to the incomplete information setting (Aumann and Maschler, 1995) are that we do not require Bayesian updating or common prior in our general set up.

In the second part of the paper we study the implications of limited awareness for equilibrium play in such games and show that even a "small amount" of unawareness is enough to yield different paths of play in equilibrium. In particular we establish conditions that ensure that awareness converges and study the restrictions on the type space needed to guarantee that the paths induced in Nash equilibria and subgame perfect Nash equilibria of the game with common knowledge of the action space can still be induced in a subjective equilibrium of the game with limited awareness. We find that first-order mutual knowledge of the action set is a necessary and sufficient condition to guarantee that every Nash equilibrium from the common knowledge case can be recovered under limited awareness in any game. For subgame perfection second order mutual knowledge is both necessary and sufficient. We also study the set of equilibrium payoffs in games where the Nash folk theorem for finitely repeated games applies (Benoit and Krishna, 1987) and show that it relates in a non-monotonic way to the awareness of agents. This arises because our framework (due to the fact that there may be unforeseen contingencies) produces a distinction between subjectively and objectively individually rational behavior in a natural way. This cannot occur in the standard theory of repeated games with incomplete information.[4]

The work that is most closely related to ours is probably the work by Halpern and Rego (2006), Sadzik (2006) or Heifetz, Meier and Schipper (2007, 2008, 2009) on the other hand. Heifetz, Meier and Schipper (2007) as well as Sadzik (2006) define Bayesian equilibrium for static games with unawareness. Halpern and Rego (2006) provide a semantic model for games with unawareness, define a Nash equilibrium for such games and show its existence. Heifetz, Meier and Schipper (2008) also provide a semantic (state space) construction for interactive unawareness. Their framework (just as Halpern and Rego, 2006) allows agents to reason about the awareness of others. They show that this construction retains a number of desired properties of unawareness when unawareness is defined as not knowing and not knowing that you don't know. Using this model they show that mutual unawareness can lead to speculative trade. In Heifetz, Meier and Schipper (2009) they extend extensive form rationalizability to this setting. Since we aim to model limited awareness as a game with incomplete information on our dynamic type space we use standard rationalizability.

Also related are Grant and Quiggin (2007) already mentioned above who discuss the notion of discovery. Feinberg (2004) provides a syntactic model to model limited awareness. He shows that in finitely repeated prisoner's dilemma games introducing a small amount of unawareness can have the same effect as the introduction of irrational types as in the seminal work by Kreps and Wilson (1982).[5] Another main difference between our approach and most of the existing

[4]See e.g. the textbook by Aumann and Maschler (1995).

[5]Other standard references include Fagin and Halpern (1988), Modica and Rustichini (1994, 1999) or Halpern

literature is that we use a set theoretic structure (common in game theory), whereas the papers mentioned above model unawareness relying on modal logic.[6] Our paper also goes beyond much of the received literature in trying to outline precise conditions under which limited awareness will matter for equilibrium predictions in games. Note also that in our approach we require an agent to be aware of actions before he can have beliefs about them avoiding the impossibility result of Dekel, Lipman and Rustichini (1998). In the language of Halpern (1988) our paper deals with explicit rather than implicit knowledge.

The paper is organized as follows. In Section 2 we introduce the basic notation and give some preliminaries. In Section 3 we introduce the basic model and in Section 4 we show how this model of unawareness can generate interesting predictions for finitely repeated games. Proofs are relegated to an Appendix.

2. Notation and preliminaries

We present some preliminaries on Polish spaces. For further reference see Kechris (1995). A topological space $(Z, \mathcal{T})$ is called Polish if it is separable and completely metrizable. A subspace of a separable metrizable space is also separable and metrizable. Examples of Polish spaces include finite sets, $\mathbb{R}^n$ and closed subsets of Polish spaces endowed with the relative topology. The countable product of Polish spaces, endowed with the product topology, is Polish. A countable intersection of open subsets of Z is called G_δ. A subspace of a Polish space is Polish if and only if it is G_δ.

For any topological space Z let $\Delta(Z)$ denote the set of all Borel probability measures, endowed with the weak topology ($\mathcal{W}$). If Z is Polish then so is $\Delta(Z)$ (Aliprantis and Border, 1994). For some $\mu \in \Delta(Z)$ let $\Gamma(\mu)$ denote its support, i.e., the set of all points $z \in Z$ such that every $T \in \mathcal{T}$ with $z \in T$ has positive measure: $\Gamma(\mu) = \{z \in Z : z \in T \in \mathcal{T} \Rightarrow \mu(T) > 0\}$. The support is the smallest closed subset of Z with measure equal to 1. If Z is separable and metrizable, the support is unique (Parthasarathy, 1967).

2.1. Repeated games

Consider a finite normal form game $G = \langle N, S, (v_i)_{i \in N} \rangle$, where $N = \{1, 2\}$ is the set of players, $S := S_1 \times S_2$ the action space, $S_i = \{s_i^1, ..., s_i^{J_i}\}$ is player i's (finite) set of actions, with s_i being

(2001).

[6] Also Li (2008), Heifetz, Meier and Schipper (2007) and Copic and Galeotti (2007) use set theoretic structures among few others. All these discuss static games, though.

the typical element of S_i, and $v_i : S \to \mathbb{R}$ player i's payoff function. Consider i's set of (mixed) strategies $\Delta(S_i) := \{\sigma_i \in \mathbb{R}_+^{J_i} : \sum_{j=1}^{J_i} \sigma_i^j = 1\}$, with typical element $\sigma_i = (\sigma_i^1, ..., \sigma_i^{J_i})$, and let σ_i^j denote the probability that σ_i assigns to s_i^j. We define the expected payoff $v_i : \Delta(S) \to \mathbb{R}$ as usual, where $\Delta(S) := \Delta(S_1) \times \Delta(S_2)$. We call G the stage game or constituent game.

Suppose that G is played repeatedly for T periods. The realized play before time t is described by the vector of action profiles played at $\tau = 1, ..., t-1$, and is called a t-history. A t-history can be identified by the vector of action profiles that have been played during the first $t-1$ periods, i.e., h_t can be rewritten as $\big(s(1), ..., s(t-1)\big) \in S^{t-1}$, where $s(\tau) = \big((s_1(\tau), s_2(\tau)\big)$ is the realized action profile at every $\tau = 1, ..., t-1$. Let H_t be the set of t-histories, with typical element h_t. Obviously there is a bijection between H_t and S^{t-1}. Let $H_1 = \{h_1\}$ be a singleton that contains only the empty history h_1 which corresponds to "no-action till now". A $(t+k)$-history h_{t+k} is called subsequent to h_t whenever the first $t-1$ elements of h_{t+k} coincide with h_t. In this case we call h_t a sub-history of h_{t+k}. The set of h_t's subsequent histories is denoted by $H(h_t)$, and the set of h_t's subsequent $(t+k)$-histories is denoted by $H_{t+k}(h_t)$. Finally, let $H := H(h_1) = \cup_{t=1}^{T} H_t$.

At every time t, players observe some h_t, the realized one. Player i's stage strategy at time t determines what i plays after having observed any possible t-history, i.e., $\sigma_i(t) : H_t \to \Delta(S_i)$. For notation simplicity we omit the index (t): We denote what the stage strategy that $\sigma_i(t)$ specifies to be played after some $h_t \in H_t$ by $\sigma_i(h_t) \in \Delta(S_i)$. We denote player i's stage strategy space at history h_t by $\Delta(S_i(h_t))$.

Let $\Delta(\overrightarrow{S_i}(h_t)) = \times_{h_\tau \in H(h_t)} \Delta(S_i(h_\tau))$ be i's (behavioral) strategy space at the history h_t, with typical element $\overrightarrow{\sigma_i}(h_t)$. Clearly, $\overrightarrow{\sigma_i}(h_t)$ specifies a contingent plan of strategies for player i at every h_τ which is subsequent to h_t.

Player i's expected payoff[7] when $\overrightarrow{\sigma}(h_t) := (\overrightarrow{\sigma_1}(h_t), \overrightarrow{\sigma_2}(h_t))$ is played is equal to

$$u_i(\overrightarrow{\sigma}(h_t)) = v_i(\sigma(h_t)) + \sum_{\tau=t+1}^{T} \sum_{h_\tau \in H_\tau(h_t)} \beta(h_\tau | \sigma(h_{\tau-1})) v_i(\sigma(h_\tau)), \tag{1}$$

where $\beta(h_\tau | \sigma(h_{\tau-1}))$ denotes the probability to reach h_τ given that $\sigma(h_{\tau-1})$ is played at $h_{\tau-1}$.

We say that $\overrightarrow{\sigma_i}(h_t)$ is a best response to $\overrightarrow{\sigma}(h_t)$ for player i at history h_t, and we write $\overrightarrow{\sigma_i}(h_t) \in BR_i(\overrightarrow{\sigma}(h_t))$, whenever

$$\overrightarrow{\sigma_i}(h_t) \in \arg \max_{\Delta(\overrightarrow{S_i}(h_t))} u_i(\overrightarrow{\sigma_i}'(h_t), \overrightarrow{\sigma_j}(h_t)).$$

Definition 2.1. A strategy profile $\overrightarrow{\sigma}(h_t)$ is a Nash equilibrium at time h_t whenever $\overrightarrow{\sigma_i}(h_t) \in BR_i(\overrightarrow{\sigma}(h_t))$ for all $i = 1, 2$. We say that $\overrightarrow{\sigma}(h_t)$ is a subgame perfect (Nash) equilibrium at h_t

[7] Instead of the aggregate payoff, we could use the average payoff, but since the horizon is finite the two are equivalent.

whenever $\overrightarrow{\sigma}(h_{t+k})$ is a Nash equilibrium for every h_{t+k} subsequent to h_t, where $\overrightarrow{\sigma_i}(h_{t+k})$ specifies the same stage strategy as $\overrightarrow{\sigma_i}(h_t)$ to every history after h_{t+k}.

For some h_t, consider the following sequence:

$$\begin{aligned}
R_i^0(h_t) &= \Delta(\overrightarrow{S_i}(h_t)) \\
R_i^1(h_t) &= \{\overrightarrow{\sigma_i}(h_t) \in \Delta(\overrightarrow{S_i}(h_t)) : \overrightarrow{\sigma_i}(h_t) \in BR_i(\overrightarrow{\sigma_j}(h_t)); \overrightarrow{\sigma_j}(h_t) \in R_j^0(h_t)\} \\
&\vdots \\
R_i^k(h_t) &= \{\overrightarrow{\sigma_i}(h_t) \in \Delta(\overrightarrow{S_i}(h_t)) : \overrightarrow{\sigma_i}(h_t) \in BR_i(\overrightarrow{\sigma_j}(h_t)); \overrightarrow{\sigma_j}(h_t) \in R_j^{k-1}(h_t)\} \\
&\vdots
\end{aligned}$$

Definition 2.2. A strategy profile $\overrightarrow{\sigma}(h_t) = (\overrightarrow{\sigma_1}(h_t), \overrightarrow{\sigma_2}(h_t))$ is rationalizable whenever

$$\overrightarrow{\sigma_i}(h_t) \in \bigcap_{k \geq 0} R_i^k(h_t)$$

for both $i = 1, 2$.

2.2. Belief hierarchies and type spaces

We consider the standard framework introduced by Brandenburger and Dekel (1993): Let $\mathcal{A}$ denote the Polish underlying space of uncertainty. A belief hierarchy of player i describes what i believes about X, what i believes that j believes about $\mathcal{A}$, and so on. Consider the following sequence:

$$\begin{aligned}
B_0 &:= \mathcal{A} \\
B_1 &:= B_0 \times \Delta(B_0) \\
&\vdots \\
B_k &:= B_{k-1} \times \Delta(B_{k-1}) \\
&\vdots
\end{aligned}$$

A belief hierarchy is a vector $\theta_i := (\mu_0, \mu_1, ...) \in \times_{k=1}^{\infty} \Delta(B_k)$, where μ_k denotes i's k-th order beliefs. Let $\Theta_0 := \times_{k=1}^{\infty} \Delta(B_k)$ denote the set of all belief hierarchies.

The following standard coherency restriction states that i's higher order beliefs cannot contradict her own lower order beliefs: A belief hierarchy satisfies coherency, i.e., $\text{marg}_{B_{k-1}} \mu_k = \mu_{k-1}$ for all $k > 0$. Although, coherency rules out the possibility that i's beliefs contradict each other, it does not exclude hierarchies such that i believes that j's beliefs are not coherent. In order to do so, we further restrict beliefs to hierarchies that satisfy common knowledge[8] of coherency, and we denote the space of those beliefs by Θ. Then, there is a homeomorphism $g : \Theta \rightarrow \Delta(\mathcal{A} \times \Theta)$. This result was independently proven by Mertens and Zamir (1985).

A type space is a tuple $((\Theta_i)_{i \in N}, (g_i)_{i \in N})$, where Θ_i is Polish and $g_i : \Theta_i \rightarrow \Delta(\mathcal{A} \times \Theta_j)$ is continuous. The type space yields a hierarchy of beliefs for every $\theta_i \in \Theta_i$: Individual i's first order beliefs are equal to the marginal distribution of $g_i(\theta_i)$, i.e., marginalize from $\Delta(\mathcal{A} \times \Theta_j)$ to $\Delta(\mathcal{A})$. In order to obtain i's second order beliefs go from Θ_i to $\Delta(\mathcal{A} \times \Theta_j)$ via g_i, and then to $\Delta(\mathcal{A} \times \Delta(\mathcal{A} \times \Theta_i))$ via g_j and marginalize to $\Delta(\mathcal{A} \times \Delta(\mathcal{A}))$ via image measures. Continue inductively to obtain the entire hierarchy of beliefs. Harsanyi (1967-68) was the first one to introduce the concept of type spaces in order to model belief hierarchies.

Mertens and Zamir (1985), and Brandenburger and Dekel (1993) completed the analysis by proving the previous result, which implies that there is a universal type space $((\Theta)_{i \in N}, (g)_{i \in N})$, i.e., one that is both terminal and complete (Siniscalchi, 2007). That is, the universal type space yields all possible hierarchies of beliefs, and at the same time the function g is onto, implying that it does not impose any further uncertainty over the hierarchies of beliefs.

3. Awareness of the action space

3.1. Hierarchies of beliefs about awareness in the stage game

Let $\mathcal{A}_i$ be the discrete topology on S_i, and let also $\mathcal{A}$ be the set of non-empty subsets $\mathcal{A}_1 \times \mathcal{A}_2$ endowed with the discrete topology. Elements $A \in \mathcal{A}$ correspond to different awareness structures that a player may have about the stage game G.

When player i is aware of some $A \in \mathcal{A}$, she also forms beliefs about what j is aware of, beliefs about j's beliefs about what i is aware of, and so on. Player i's hierarchies of beliefs are modeled in the standard way (see Section 2.2).

Unlike most underlying spaces of uncertainty – where we are interested in the set of all possible hierarchies of beliefs – not all hierarchies over $\mathcal{A}$ are relevant. That is, given the nature of awareness, there are beliefs that simply do not make sense. Therefore, we need to place some restrictions which eliminate certain hierarchies:

(R_1) Player i knows what she is aware of, i.e., $\Gamma(\mu_0)$ is a singleton.

[8] In general, knowledge is conceptually different from belief with probability 1, but we consider them as equivalent for terminology simplicity (Brandenburger and Dekel, 1987).

Let $\Theta_1^* \subseteq \Theta_0$ contain the hierarchies that satisfy (R_1). That is, the space of first order beliefs becomes a copy $\mathcal{A}$. This restriction alone does not rule out the possibility that i believes that j does not know what she is aware of. We do that by requiring (R_1) to be commonly known, and we denote the space of hierarchies that satisfy this requirement by $\Theta_1 \subseteq \Theta_1^*$.

(R_2) Player i of type $\theta_i \in \Theta_1$ cannot believe that j is aware of an action that i herself is not aware of, i.e., if $(A, \mu_0) \in \Gamma(\mu_1)$, then $A' \subseteq A$, where $\Gamma(\mu_0) = \{A'\}$.

Let $\Theta_2^* \subseteq \Theta_1$ contain the hierarchies that satisfy $(R_1) - (R_2)$ and common knowledge of (R_1). Once again, (R_2) does not rule out the possibility that i believes that j believes that i is aware of an action that i believes that j is not aware of. We do that by requiring $(R_1) - (R_2)$ to be commonly known, and we denote the space of hierarchies that satisfy this requirement by $\Theta_2 \subseteq \Theta_2^*$. Finally, as usual, we impose the standard coherency restriction:

(R_3) A belief hierarchy satisfies coherency, i.e., $\text{marg}_{B_{k-1}} \mu_k = \mu_{k-1}$ for all $k > 0$.

Let $\Theta_3^* \subseteq \Theta_2$ contain the hierarchies that satisfy $(R_1) - (R_3)$ and common knowledge of $(R_1) - (R_2)$. Once again, (R_3) does not rule out the possibility that i believes that j believes that i is aware of an action that i believes that j is not aware of. We do that by requiring $(R_1) - (R_3)$ to be commonly known, and we denote the space of hierarchies that satisfy this requirement by $\Theta_3 \subseteq \Theta_3^*$. We call Θ_3 the universal space of awareness-consistent belief hierarchies. Then, we extend the standard result to this setting that accounts for awareness-consistency.

Proposition 3.1. *There is a homeomorphism $g_3 : \Theta_3 \to \Delta(\Theta_3)$.*

The previous result implies that Θ_3 is both terminal and complete. That is, all belief hierarchies that satisfy $(R_1) - (R_3)$ and common knowledge of $(R_1) - (R_3)$ belong to Θ_3, and also the function g_3 is onto (see again Section 2.2).

3.2. Awareness in repeated games

In the previous section we lay the foundations for modeling hierarchies of beliefs about awareness in the stage game. However, that setting is entirely static, which may not allow to capture certain states of mind. To see this consider the following simple example.

Example 3.1. Suppose that there are two different contingencies (time periods), $t = 1, 2$, and a symmetric normal form game with $S = \{a, b, c\}$. At $t = 1$ player i:

- is aware of actions a and b,

- believes that at $t = 1$ player j is aware only of action a,
- believes that at $t = 2$ player j will be aware of a and b.

In principle there is no reason why we should assume that i believes that j's awareness remains constant over time. Thus, at $t = 1$ player i forms beliefs about j's awareness at every t, also forms beliefs about j's beliefs at every t about i's awareness at every t, and so on.

At the same time, i's awareness and beliefs at $t = 2$ need not be the same as at $t = 1$, but still i at $t = 1$ cannot foresee these changes, thus believing at $t = 1$ that in the future neither her own awareness nor her beliefs about j's awareness at every contingency will change. $\triangleleft$

In a repeated game setting the different contingencies correspond to the different histories. In order to capture i's entire hierarchy of beliefs at some history h we consider the following sequence:

$$\begin{aligned}
\overrightarrow{B_0} &:= \mathcal{A} \\
\overrightarrow{B_1} &:= \overrightarrow{B_0} \times \underbrace{\Delta(\overrightarrow{B_0}) \times \cdots \times \Delta(\overrightarrow{B_0})}_{h \in H} = \overrightarrow{B_0} \times \left(\times_{h \in H} \Delta(\overrightarrow{B_0})\right) \\
&\vdots \\
\overrightarrow{B_k} &:= \overrightarrow{B_{k-1}} \times \underbrace{\Delta(\overrightarrow{B_{k-1}}) \times \cdots \times \Delta(\overrightarrow{B_{k-1}})}_{h \in H} = \overrightarrow{B_{k-1}} \times \left(\times_{h \in H} \Delta(\overrightarrow{B_{k-1}})\right) \\
&\vdots
\end{aligned}$$

A belief hierarchy after h_t is a vector $\overrightarrow{\theta_i}(h_t) := (\overrightarrow{\mu_0}, \overrightarrow{\mu_1}, ...) \in \times_{k=1}^{\infty} \Delta(\overrightarrow{B_k})$, where $\overrightarrow{\mu_k}$, denotes i's k-th order beliefs. Let $\overrightarrow{\Theta_0} := \times_{k=1}^{\infty} \Delta(\overrightarrow{B_k})$ denote the set of all belief hierarchies that i may hold at some h.

Of course it is possible that player i is not aware of the existence of some histories, i.e., the ones that in order to be reached an action that i is not aware of needs to be played. Thus, similar to the static case, we impose a number of restrictions on the belief hierarchies, which arise as a consequence of the nature of awareness.

(R_1') Player i knows what she is aware of, i.e., $\Gamma(\overrightarrow{\mu_0})$ is a singleton.

Definition 3.1. Let i's awareness given some type $\overrightarrow{\theta_i}$ be denoted by $a(\overrightarrow{\theta_i}) := a_1(\overrightarrow{\theta_i}) \times a_2(\overrightarrow{\theta_i})$, where $a_j(\overrightarrow{\theta_i})$ denotes the set of j's actions that i is aware of for each $j = 1, 2$.

That is, if $\Gamma(\overrightarrow{\mu_0}) = \{A\}$ then $a(\overrightarrow{\theta_i}) = A$. Let $\overrightarrow{\Theta_1}^* \subseteq \overrightarrow{\Theta_0}$ contain the hierarchies that satisfy (R_1'), and $\overrightarrow{\Theta_1} \subseteq \overrightarrow{\Theta_1}^*$ contain those hierarchies that satisfy (R_1') and common knowledge of (R_1').

(R_2') At every history $h_t = \big(s(1), ..., s(t-1))\big) \in S^{t-1}$ player i is aware of all action that have been played before reaching this history, i.e., player i's type $\overrightarrow{\theta_i}$ satisfies $\{s_1(\tau)\} \times \{s_2(\tau)\} \subseteq a(\overrightarrow{\theta_i})$ for all $\tau = 1, ..., t-1$.

Again, let $\overrightarrow{\Theta_2}^* \subseteq \overrightarrow{\Theta_1}$ contain the hierarchies that satisfy $(R_1') - (R_2')$ and common knowledge of (R_1'), and $\overrightarrow{\Theta_2} \subseteq \overrightarrow{\Theta_2}^*$ contain those hierarchies that also satisfy common knowledge of (R_2').

(R_3') Player i of type $\overrightarrow{\theta_i} \in \overrightarrow{\Theta_2}$ cannot believe that after any history j is aware of an action that i herself is not aware of, i.e., if $\big(A, (\overrightarrow{\mu_0}^h; h \in H)\big) \in \Gamma(\overrightarrow{\mu_1})$, then $A_h \subseteq A$ for every $h \in H$, where $\Gamma(\overrightarrow{\mu_0}^h) = \{A_h\}$.

Again, let $\overrightarrow{\Theta_3}^* \subseteq \overrightarrow{\Theta_2}$ contain the hierarchies that satisfy $(R_1') - (R_3')$ and common knowledge of $(R_1') - (R_2')$, and $\overrightarrow{\Theta_3} \subseteq \overrightarrow{\Theta_3}^*$ contain those hierarchies that satisfy $(R_1') - (R_3')$ and common knowledge of $(R_1') - (R_3')$.

(R_4') Player i of type $\overrightarrow{\theta_i} \in \overrightarrow{\Theta_3}$ believes that j has perfect recall, i.e., if $\big(A, (\overrightarrow{\mu_0}^h; h \in H)\big) \in \Gamma(\overrightarrow{\mu_1})$ and h is a sub-history of h', then $A_h \subseteq A_{h'}$, where $\Gamma(\overrightarrow{\mu_0}^h) = \{A_h\}$ and $\Gamma(\overrightarrow{\mu_0}^{h'}) = \{A_{h'}\}$.

Once again, let $\overrightarrow{\Theta_4}^* \subseteq \overrightarrow{\Theta_3}$ contain the hierarchies that satisfy $(R_1') - (R_4')$ and common knowledge of $(R_1') - (R_3')$, and $\overrightarrow{\Theta_4} \subseteq \overrightarrow{\Theta_4}^*$ contain those hierarchies that satisfy $(R_1') - (R_4')$ and common knowledge of $(R_1') - (R_4')$.

Finally, we impose the standard coherency requirement:

(R_5') A belief hierarchy satisfies coherency, i.e., $\text{marg}_{\overrightarrow{B_{k-1}}} \overrightarrow{\mu_k} = \overrightarrow{\mu_{k-1}}$ for all $k > 0$.

Let $\overrightarrow{\Theta_5}^* \subseteq \overrightarrow{\Theta_4}$ denote the space of coherent belief hierarchies, whilst $\overrightarrow{\Theta_5} \subseteq \overrightarrow{\Theta_5}^*$ denotes the space of hierarchies that satisfy $(R_1') - (R_5')$ and common knowledge of $(R_1') - (R_5')$. We call $\overrightarrow{\Theta_5}$ the universal awareness-consistent type space for the repeated game.

Proposition 3.2. *There is a homeomorphism* $\overrightarrow{g_5} : \overrightarrow{\Theta_5} \to \Delta(\times_{h \in H} \overrightarrow{\Theta_5})$.

Though an element of $\overrightarrow{\Theta_5}$ refers to i's beliefs about the awareness structure at other than the current history, it is still static in the sense that it does not induce any beliefs about i's own awareness and beliefs at other histories. However, in this particular setting the nature of awareness is such that introducing further uncertainty about the own type $\overrightarrow{\theta_i}$ at other histories would induce a redundancy. The reason is that i at a history h is confident that at every subsequent history she will have the same type, i.e., she is unaware of the possibility that her current awareness and beliefs may be wrong. At the same type she has perfect recall in the sense that she remembers her past types, and therefore she assigns probability 1 to them. We will further discuss this matter in more details in the following section.

Definition 3.2. A type space is a tuple $((\overrightarrow{\Theta_i})_{i\in N}, (\overrightarrow{g_i})_{i\in N})$, where $\overrightarrow{\Theta_i} \subseteq \overrightarrow{\Theta_5}$ is Polish, and $\overrightarrow{g_i}$: $\overrightarrow{\Theta_i} \to \Delta(\times_{h\in H}\overrightarrow{\Theta_j})$ is continuous and agrees with $\overrightarrow{g_5}$ on $\overrightarrow{\Theta_i}$.

For notation simplicity, let $\overrightarrow{g_i}(\cdot|\overrightarrow{\theta_i}) := \overrightarrow{g_i}(\overrightarrow{\theta_i})$. Let also $\text{marg}_h\, \overrightarrow{g_i}(\overrightarrow{\theta_i}) \in \Delta(\overrightarrow{\Theta_j})$ denote the marginal of $\overrightarrow{g_i}(\overrightarrow{\theta_i})$ that corresponds to the type of the opponent at history h, i.e., it determines i's current beliefs about j's type at the history h.

Definition 3.3. A behavioral strategy for player i at h in a repeated game with limited awareness is a function $\overrightarrow{\sigma_i}(h) : \overrightarrow{\Theta_i} \to \Delta(\overrightarrow{S_i}(h))$.

The (repeated game) strategy that $\overrightarrow{\sigma_i}(h)$ induces for a particular type $\overrightarrow{\theta_i}$ is denoted by $\overrightarrow{\sigma_i}(h, \overrightarrow{\theta_i})$, and the stage game strategy that $\overrightarrow{\sigma_i}(h, \overrightarrow{\theta_i})$ induces is denoted by $\sigma_i(h, \overrightarrow{\theta_i})$.

Remark 3.1. Player i's strategy determines what i's contingent plan at h would be for every possible type $\overrightarrow{\theta_i}$. Of course, in reality at h player i of type $\overrightarrow{\theta_i}$ is aware of only the histories that do not involve any action she is not aware of (see Restriction (R_2')). Therefore, she cannot really make a contingent plan for those histories that she is not aware of. However, for notation simplicity we assume that $\overrightarrow{\sigma_i}(h)$ determines a behavioral strategy all types and all subsequent histories. This does not affect neither our results nor our intuition, since the behavioral strategies that correspond to these histories do not enter i's hierarchy of beliefs. ◁

The expected payoff to i at h_t given $\overrightarrow{\theta_i}$ is equal to

$$\begin{aligned} u_i(\overrightarrow{\sigma}(h_t)|\overrightarrow{\theta_i}) &= \sum_{\overrightarrow{\theta_j}\in\overrightarrow{\Theta_j}} \text{marg}_{h_t}\, \overrightarrow{g_i}(\overrightarrow{\theta_j}|\overrightarrow{\theta_i}) v_i(\sigma_i(h_t, \overrightarrow{\theta_i}), \sigma_j(h_t, \overrightarrow{\theta_j})) \\ &+ \sum_{\tau=t+1}^{T} \sum_{h_\tau\in H_\tau(h_t)} \beta(h_\tau|\overrightarrow{\sigma}(h_{\tau-1})) \sum_{\overrightarrow{\theta_j}\in\overrightarrow{\Theta_j}} \text{marg}_{h_\tau}\, \overrightarrow{g_i}(\overrightarrow{\theta_j}|\overrightarrow{\theta_i}) v_i(\sigma_i(h_\tau, \overrightarrow{\theta_i}), \sigma_j(h_\tau, \overrightarrow{\theta_j})). \end{aligned}$$

We say that $\overrightarrow{\sigma_i}(h_t)$ is a best response to $\overrightarrow{\sigma}(h_t)$ for player i given $\overrightarrow{\theta_i}$ at h_t, and we write $\overrightarrow{\sigma_i}(h_t, \overrightarrow{\theta_i}) \in BR_i(\overrightarrow{\sigma_j}(h_t)|\overrightarrow{\theta_i})$, whenever

$$\overrightarrow{\sigma_i}(h_t, \overrightarrow{\theta_i}) \in \arg\max_{\Delta(\overrightarrow{S_i}(h_t))} u_i(\overrightarrow{\sigma_i}'(h_t), \overrightarrow{\sigma_j}(h_t)|\overrightarrow{\theta_i}).$$

We say that $\overrightarrow{\sigma_i}(h_t)$ is a best response to $\overrightarrow{\sigma}(h_t)$ for player i at h_t, and we write $\overrightarrow{\sigma_i}(h_t) \in BR_i(\overrightarrow{\sigma_j}(h_t))$, whenever $\overrightarrow{\sigma_i}(h_t)$ is a best response to $\overrightarrow{\sigma}(h_t)$ for player i given every $\overrightarrow{\theta_i} \in \overrightarrow{\Theta_i}$ at h_t.

3.3. Equilibrium with limited awareness

The equilibrium concept that we propose has a similar flavor as the standard Bayesian Nash equilibrium (Harsanyi, 1967-68). Though, the players may have wrong beliefs about the awareness structure of the opponent, in equilibrium they have correct beliefs about what the opponent does given every awareness structure that they are aware of, i.e., i has correct beliefs about the j's behavioral strategy given each type $\overrightarrow{\theta_j}$. Of course, i in principle forms beliefs only about what j does at the types $\overrightarrow{\theta_j} \in \Gamma(\overrightarrow{g_i}(\overrightarrow{\theta_i})$ that i considers as possible. Therefore, what j does at other types is in any case completely irrelevant from i's point of view, since i assigns probability 0 to those types that she is not aware of. Formally:

Definition 3.4. A strategy profile $\overrightarrow{\sigma}(h_t) = (\overrightarrow{\sigma_1}(h_t), \overrightarrow{\sigma_2}(h_t))$ is a Nash equilibrium at h_t in the repeated game with limited awareness, whenever

$$\overrightarrow{\sigma_i}(h_t, \overrightarrow{\theta_i}) \in BR_i(\overrightarrow{\sigma}(h_t)|\overrightarrow{\theta_i})$$

for all $\overrightarrow{\theta_i} \in \overrightarrow{\Theta_i}$ and for all $i = 1, 2$.

Note that we define Nash equilibrium explicitly for all histories h_t since we have to take into account the possibility that awareness of either or both agents might change during the course of play.

Proposition 3.3. *A Nash equilibrium exists in a finite horizon repeated game with limited awareness.*

Definition 3.5. A strategy profile $\overrightarrow{\sigma}(h_t) = (\overrightarrow{\sigma_1}(h_t), \overrightarrow{\sigma_2}(h_t))$ is a subgame perfect Nash equilibrium in the repeated game with limited awareness, whenever $\overrightarrow{\sigma}(h_{t+k})$ is a Nash equilibrium at every $h_{t+k} \in H(h_t)$, where $\overrightarrow{\sigma}(h_{t+k})$ specifies the same strategy as $\overrightarrow{\sigma}(h_t)$ to every history in $H(h_t)$.

Intuitively though, the definition of (subgame perfect) Nash equilibrium is sometimes requiring too much. The reason is that Nash equilibrium requires agents to have correct beliefs about the action choices of all types they are aware of - even those that are "off equilibrium path". In particular, since we assume that players are confident about their type we also want to explore an equilibrium concept which requires less reasoning about the opponent's type. The alternative definition of "subjective Nash equilibrium" that we present below requires only that agent i 's beliefs about j's action choices should not be contradicted by optimal behavior on the equilibrium path.

In order to do this we need some more notation. Denote by $\xi_i(\overrightarrow{\theta_j}, h_t|\overrightarrow{\theta_i})$ the belief player i of type $\overrightarrow{\theta_i}$ has about type $\overrightarrow{\theta_j}$'s action choice at history h_t and let $\xi_i(h_t|\overrightarrow{\theta_i})$ be the total probability agent i attaches to j's action choices given his type $\overrightarrow{\theta_i}$ (i.e. the componentwise product of $\xi_i(\overrightarrow{\theta_j}, h_t|\overrightarrow{\theta_i})$ and $\text{marg}_{h_t}\, \overrightarrow{g_i}(\overrightarrow{\theta_j}|\overrightarrow{\theta_i})$). To save notation we will sometimes denote the latter simply by $\xi_i(h_t)$. Furthermore we denote a players realized (ex post) type by $\widehat{\theta}_{-i}$.

Definition 3.6. A strategy profile $\overrightarrow{\sigma}(h_t) = (\overrightarrow{\sigma_1}(h_t), \overrightarrow{\sigma_2}(h_t))$ is a subjective Nash equilibrium at h_t in the repeated game with limited awareness, whenever

$$\begin{array}{rcl} \text{(i) } \overrightarrow{\sigma_i}(h_t, \overrightarrow{\theta_i}) & \in & BR_i(\xi_i(h_t)|\overrightarrow{\theta_i}) \\ \text{(ii) } \overrightarrow{\sigma}_{-i}(h_t, \widehat{\theta}_{-i}) & \in & \Gamma(\xi(h_t)) \end{array}$$

for all $\overrightarrow{\theta_i} \in \overrightarrow{\Theta_i}$ and for all $i = 1, 2$.

The definition of subjective subgame perfect Nash equilibrium follows in the obvious way.

3.4. Change of types over time

Although player i is certain about her past type and confident about the future type, this does not mean that her type will remain constant throughout all histories. There are two questions that arise at this point:

- Which are the types that i may have given the types that she had at every sub-history?
- Under what conditions does a player's type change over time?

We start from the first question: Given that i remembers at h what her type was at all preceding histories, the only requirement is the following:

(R_6') For two histories $h, h' \in H$, with h being subsequent to h', if $\overrightarrow{\theta_i}$ and $\overrightarrow{\theta_i}'$ are the corresponding types, then $a(\overrightarrow{\theta_i}) \subseteq a(\overrightarrow{\theta_i}')$.

Now, we turn to the second question. In principle, at every h_t player i is confident that her awareness and beliefs described by $\overrightarrow{\theta_i}$ are correct, and therefore unless she observes something unexpected her type will remain the same. Therefore, we need to define what "unexpected observation" means. In order to do that, it is necessary to define what is a $\overrightarrow{\theta_i}$-rationalizable strategy at h_t. Let $\Delta(\overrightarrow{a_i}(\overrightarrow{\theta_i}, h_t)) := \times_{h \in H(h_t)} \Delta(a_i(\overrightarrow{\theta_i}))$ be the action space that i is aware of h_t when her type is $\overrightarrow{\theta_i}$, and consider the following sequence:

$$\begin{array}{rcl} \overrightarrow{R_i}^0(h_t) & = & \{\overrightarrow{\sigma_i}(h_t) \in [\Delta(\overrightarrow{S_i}(h_t))]^{\overrightarrow{\Theta_i}} : \overrightarrow{\sigma_i}(h_t, \overrightarrow{\theta_i}') \in \Delta(\overrightarrow{a_i}(\overrightarrow{\theta_i}'))\} \\ \overrightarrow{R_i}^1(h_t) & = & \{\overrightarrow{\sigma_i}(h_t) \in R_i^0(h_t) : \overrightarrow{\sigma_i}(h_t) \in BR_i(\overrightarrow{\sigma_j}(h_t)); \overrightarrow{\sigma_j}(h_t) \in \overrightarrow{R_j}^0(h_t)\} \\ & \vdots & \\ \overrightarrow{R_i}^k(h_t) & = & \{\overrightarrow{\sigma_i}(h_t) \in R_i^0(h_t) : \overrightarrow{\sigma_i}(h_t) \in BR_i(\overrightarrow{\sigma_j}(h_t)); \overrightarrow{\sigma_j}(h_t) \in \overrightarrow{R_j}^{k-1}(h_t)\} \\ & \vdots & \end{array}$$

Definition 3.7. We say that a strategy profile $\overrightarrow{\sigma}(h_t) = (\overrightarrow{\sigma_1}(h_t), \overrightarrow{\sigma_2}(h_t))$ is rationalizable whenever

$$\overrightarrow{\sigma_i}(h_t) \in \bigcap_{k \geq 0} \overrightarrow{R_i}^k(h_t),$$

for both $i = 1, 2$.

Definition 3.8. We say that a behavioral strategy $\overrightarrow{\sigma_i}(h_t, \overrightarrow{\theta_i})$ is $\overrightarrow{\theta_i}$-rationalizable if it is the $\overrightarrow{\theta_i}$ element of a rationalizable strategy profile $\overrightarrow{\sigma}(h_t)$.

The set of $\overrightarrow{\theta_i}$-rationalizable strategies contains those contingent plans that i believes that they can be rationalized given her awareness structure. For some $\overrightarrow{\theta_j} \in \Gamma(\text{marg}_{h_t} \overrightarrow{g_i}(\overrightarrow{\theta_i}))$ the set of $\overrightarrow{\theta_j}$-rationalizable strategies contains those behavioral strategies that i believes that j will be able to rationalize in case j's type is $\overrightarrow{\theta_j}$, which is something that i considers as possible. Because of restriction (R_3') player i believes that she must be able to rationalize more strategies than j.

Proposition 3.4. *If a strategy profile is $\overrightarrow{\theta_j}$-rationalizable at h_t then it is also $\overrightarrow{\theta_i}$-rationalizable for every $\overrightarrow{\theta_j} \in \Gamma(\text{marg}_{h_t} \overrightarrow{g_i}(\overrightarrow{\theta_i}))$.*

Remark 3.2. The converse is not necessarily true, i.e., i may believe that a strategy is rationalizable given her type $\overrightarrow{\theta_i}$ and at the same time believe that j believes that it is not, given the type $\overrightarrow{\theta_j}$ which i considers as possible. $\triangleleft$

Assumption 3.1. Every player plays a $\overrightarrow{\theta_i}$-rationalizable strategy at every history and for every $\overrightarrow{\theta_i}$ and this is commonly known. $\triangleleft$

4. Results with limited awareness

The aim of this section is twofold. On the one hand we would like to point out the conditions under which awareness converges. The second set of results refer to the induced equilibrium path in any (subgame perfect) Nash equilibrium of the repeated game with limited awareness, where roughly speaking we show that a "small amount of unawareness" is enough to yield outcomes that are very different from the standard common knowledge case. We will also show when the standard folk theorems apply and when they do not in games with limited awareness.

4.1. Convergence of Awareness

A natural question given our setting is whether there are conditions that ensure that agents will not learn about new actions anymore, i.e. whether there exists a time t^* s.t. $\forall t \geq t^*$ the set $a(\overrightarrow{\theta_i}, h_t)$ stays constant. More precisely let's use the following definition.

Definition 4.1. We say that awareness has converged at time t^* whenever $\forall t \geq t^* : a(\overrightarrow{\theta_i}, h_{t^*}) = a(\overrightarrow{\theta_i}, h_t)$.

Ideally one would like to be able also to say something about the convergence of types $\overrightarrow{\theta}$. Since (except for $(R_1') - (R_5')$) we do not impose restrictions on higher order beliefs, we can never ensure pointwise convergence of types. For the same reason also convergence to some invariant subset of $\overrightarrow{\Theta}$ can only be ensured by imposing trivial conditions. This is why for now we focus on convergence of action awareness. In order to state the following proposition we still need a little more notation. Denote by S^{NE} the support of the set of Nash equilibria of the repeated game. Then we can state the following proposition

Proposition 4.1. *Generally awareness converges at t^*whenever either (i) $a^* := a(\overrightarrow{\theta_1}, h_{t^*}) = a(\overrightarrow{\theta_2}, h_{t^*})$ or (ii) S^{NE} is first-order mutual knowledge. Trivial convergence occurs if $a^* = S$.*

The intuition for the Proposition is quite obvious. Note that the only way to learn about new actions is either to observe them or to find them through exploration after observing something non rationalizable. If both players know exactly the same set of actions and if this fact is first-order mutual knowledge neither condition can ever obtain. The same is true if S^{NE} is first order mutual knowledge, as can be seen in the Appendix.

4.2. The Induced Equilibrium Path

4.2.1. (Subgame Perfect) Nash Equilibria

In this subsection we establish necessary and sufficient conditions on the type space $\overrightarrow{\Theta}$ that ensure that an equilibrium in the game with common knowledge continues to be an equilibrium in the game with unawareness given the restricted type space. We will employ the following definition.

Definition 4.2. Denote by $\overrightarrow{\Theta}^k \subseteq \overrightarrow{\Theta}$ the type space obtained from $\overrightarrow{\Theta}$ by making the additional restriction that the action set S is at least k-th order mutual knowledge $\forall \overrightarrow{\theta} \in \overrightarrow{\Theta}^k$.

Obviously the space $\overrightarrow{\Theta}^\infty$ corresponds to the standard case where the action set is common knowledge among the players. Note also that for any k it is the case that $\overrightarrow{\Theta}^{k+1} \subset \overrightarrow{\Theta}^k$. The following Proposition shows that the conditions on $\overrightarrow{\Theta}$ needed for a Nash equilibrium (NE) path from the game with common knowledge to be induced even with some unawareness are fairly week. Denote by $\overrightarrow{\sigma}^{NE\infty}$ a Nash equilibrium of the finitely repeated game with type space $\overrightarrow{\Theta}^\infty$.

Proposition 4.2. *We can be sure that for any stage game G and any $\overrightarrow{\sigma}^{NE\infty}$ there is a subjective NE $\forall \widehat{\theta} \in \overrightarrow{\Theta}^k$that induces the path induced by $\overrightarrow{\sigma}^{NE\infty}$if and only if $k \geq 1$.*

Hence, if and only if we restrict the type space in such a way that the action set is (at least) first-order mutual knowledge we can retrieve any NE-path of the original game in a subjective NE of the game with unawareness. The intuition is quite obvious. Assume that players do not have first order mutual knowledge of S. Then if both players believe that their opponent is only aware of $S'_j \subset S_j$ they will choose a best response to beliefs contained in $\Delta(S'_j)$ what need obviously not induce a NE path. Note that the qualifier "any stage game" is important here, since it is obviously the case that games can be found where first order-mutual knowledge is not needed to maintain a NE path. On the other hand if the action set is first-order mutual knowledge then agents beliefs will be on $\Delta(S_j)$ and hence all Nash equilibria can be maintained.[9]

Proposition 4.3. *We can be sure that for any stage game G and any $\overrightarrow{\sigma}^{SPNE\infty}$ there is a subjective SPNE $\forall \widehat{\theta} \in \overrightarrow{\Theta}^k$ that induces the path induced by $\overrightarrow{\sigma}^{SPNE\infty}$ if and only if $k \geq 2$.*

The intuition for this result is slightly more complicated and best explained with an example. Consider the following game

	X	Y	Z
A	3,2	2,1	1,3
B	4,8	12,13	**7,15**
C	2,17	18,18	6,23

which we assume is first-order mutual knowledge. (Note that since this stage game has a unique Nash equilibrium, the unique $\overrightarrow{\sigma}^{SPNE\infty}$ involves players choosing (B, Z) at each period). Suppose now, that S is *not* second-order mutual knowledge. More specifically, assume that the column player (CP) believes that the row player (RP) believes that the CP does not know action Z. (Formally $\mathrm{marg}_{h_1} \overrightarrow{g_{CP}}(\overrightarrow{\theta}_{k_1}|\widehat{\theta}_{CP}) = 1$ and $\mathrm{marg}_{h_1} \overrightarrow{g_{CP}}(\overrightarrow{\theta_{k_2}}|\overrightarrow{\theta_{k_1}}) = 1$, where $a(\overrightarrow{\theta_{k_1}}, h_t) = S$ but $a(\overrightarrow{\theta_{k_2}}, h_t) = \{A, B, C\} \times \{X, Y\}$). On the other hand assume that the RP knows the game and knows the CP's true type. The strategy profile inducing $(B, Z), \forall t$ is not a (subjective) SPNE given these types, since in all subgames following histories h_t not containing Z, the CP will expect the RP to choose C (since the unique subjective NE in the one shot game defined through the restriction to $S' = \{A, B, C\} \times \{X, Y\}$ is given by (C, Y)). In subgames following histories containing Z the CP will expect the RP to choose B[10]. But then the unique best response of the CP is to choose

[9]Note also the relation to the result by Aumann and Brandenburger (1995).

[10]Note though, that choosing (B, Z),$\forall t$ is a subjective NE (by Proposition 4.2). The reason is that if the CP holds the (subjectively wrong) off equilibrium belief that the RP will choose B also after histories not containing Z his best response is to choose Z in all periods hence such subgames will not be reached.

Y after histories that do not contain Z. The following path is induced in a SPNE

$$\underbrace{(C,Y) \to (C,Y)... \to (C,Y)}_{t=1,..T-2} \to (B,Z) \to (B,Z).$$

Since the CP believes that the RP believes that the CP does not know Z, he will choose Y in periods $t = 1,..T-2$ in order not to reveal his knowledge of Z to the row player. The RP can see through this "deception" but since it is in her interest not to reveal her type she will choose C. At $T-1$, though, the CP will choose Z in order to reap the deviation payoff of 23. Since the RP anticipates this she will choose B at $T-1$. This means that the CP realizes he was holding a wrong assessment of the RP's knowledge and updates his type accordingly. Observe that as $T \to \infty$ the outcome will be (C,Y) almost all the time in this equilibrium.

Note also that if S were second order mutual knowledge this could not be an equilibrium path, since then at $T-2$ the CP (knowing that the RP knows that he knows Z) would expect the RP to choose B at $T-1$ and hence would have no incentives to choose Y himself at $T-1$. Finally note that in games (such as the one above), where there is no Nash folk theorem, second-order mutual knowledge is necessary only if the SPNE in question is not pareto-efficient. The reason simply is that whenever the SPNE of the game with common knowledge is pareto efficient at least one player will not have an incentives to avoid the "subjective subgames" where one of the SPNE actions is not common knowledge. In the next subsection we will analyze finitely repeated games where we have folk theorems.

Remark 4.1. *Note that both Propositions 4.2 and 4.3 also hold in a Nash equilibrium for some realized types with first (second) - order mutual knowledge. The sufficiency part will not hold for all types in* $\overrightarrow{\Theta}^1$ *(*$\overrightarrow{\Theta}^2$*), though.*

4.2.2. Folk Theorems

In this subsection we consider stage games G that satisfy the conditions for the Nash folk theorem for finitely repeated games (Benoit and Krishna, 1985) and ask which payoff vectors can be sustained in a Nash equilibrium as we put more and more restrictions on the type space $\overrightarrow{\Theta}$.

Obviously without restrictions on $\overrightarrow{\Theta}$ (other than $(R_1') - (R_5')$) any game payoff can be induced at a Nash equilibrium by simply choosing the appropriate types.[11] On the other hand under the most extreme restriction on the type space, where we require common knowledge of the entire game ($\overrightarrow{\Theta}^{\infty}$), we obviously have the folk theorem by assumption, i.e. know that every individually rational (IR) payoff vector can be approximated at a Nash equilibrium. Interestingly it is not

[11] Note that one can always choose types s.t. each player is only aware of one particular action.

true that by placing more restrictions on $\overrightarrow{\Theta}$ we monotonically get sharper predictions. This is illustrated by the following proposition.

Proposition 4.4. *Consider any game G with Nash folk theorem and assume the relevant space is $\overrightarrow{\Theta}^0$. Then it is neither true (in general) that payoff vectors which are not IR can be ruled out at a Nash equilibrium of the repeated game based on G, nor is it true that every IR payoff vector can be approximated.*

If both players know the entire action set (but do not necessarily know that the other player does so), then the equilibrium prediction of the game with limited awareness is neither nested by nor does it nest the prediction of the case with common knowledge. Note that in the Proposition we use individual rationality as an objective concept taken from the case of common knowledge. Obviously, since we assume that players - given their awareness - are fully rational their behavior will always be "subjectively individually rational".

The intuition behind this Proposition is again best illustrated with an example. Let's start with the first part, which is less obvious. Consider the following stage game, which we assume is repeated for $T = 4$ periods.

	X	Y	Z
A	**5,6**	0,5	5,5
B	-50,-12	-5,-11	**100,-10**
C	3,2	**2,5**	3,5

We assume that both players know the entire action set, (i.e. restrict to $\overrightarrow{\Theta}^0$) but we assume that the row player (RP) has the following beliefs about the column player's type at the beginning of the game (at $t = 1$). He attaches high probability to the column player (CP) being of type $\overrightarrow{\theta_Y}$ where $a(\overrightarrow{\theta_Y}) = \{A, B, C\} \times \{Y\}$ and low probability to the CP being of type $\overrightarrow{\theta_X}$ where $a(\overrightarrow{\theta_X}) = \{A, B, C\} \times \{X\}$. Say $\text{marg}_{h_1} \overrightarrow{g_{RP}}(\overrightarrow{\theta_Y}|\overrightarrow{\theta_{RP}}) = 0.99$ and $\text{marg}_{h_1} \overrightarrow{g_{RP}}(\overrightarrow{\theta_X}|\overrightarrow{\theta_{RP}}) = 0.01$. Assume also that the CP knows the type of the RP, i.e., knows that he is holding this first-order belief and that Bayesian updating of beliefs is common knowledge. Then the following can be an induced equilibrium path:

$$(B, X) \to (A, X) \to (A, X) \to (A, X).$$

On this path then the payoff vector that the RP obtains is clearly not individually rational. Why can this be an equilibrium ? Note first that the RP would obviously like to reach an equilibrium where (B, Z) is played along the equ. path. Now given his type at $t = 1$ the only way he sees to achieve this is to choose an action that he believes the CP will find non rationalizable and hope that as a consequence the CP will "learn" about Z (with some probability ε). Whenever ε is "big enough" he will find it optimal to choose B. Now consider the decision of the CP. She knows the type of the RP and hence she can anticipate his reasoning in equilibrium. But given this it will be optimal for her to choose X in order to make the row player belief that she only knows X. Note that the RP in turn cannot anticipate this reasoning, since he is strictly less aware than the CP. Note that in a standard repeated game setting with incomplete information such a result could not be obtained since in this setting every subjectively individually rational payoff vector is also objectively individually rational.

The second part of the Proposition is also illustrated easily using the game above. Assume e.g. that the RP believes that the CP does not know Y. Then the individually rational payoff vector of $(2, 5)$ cannot be approximated.

Some more remarks are at order. First one might wonder what the lowest possible sustainable payoff vector is given type space $\overrightarrow{\Theta}^0$? In general this is simply the second-lowest payoff in the game. Also note that as soon as the action set is first-order mutual knowledge we obtain again the Nash folk theorem, which follows essentially from Proposition 4.2. Finally note that payoff vectors which are not (objectively) individually rational can only be part of an equilibrium (if the action set is mutual knowledge) if awareness converges strictly after period $t = 1$. In the next subsection we want to discuss the implications of observing non-rationalizable behavior on the equilibrium path a little more closely.

4.2.3. Non-rationalizable Behavior

One of the questions we want to ask in this subsection is which conditions on the initial types have to be satisfied s.t. the set of sustainable payoffs that the agents foresee at $t = 1$ coincides with those that are actually sustainable given their types ? The answer to this question will also delimit the range of situations where our model can be thought of as a standard model of repeated games with incomplete information, since if these conditions are satisfied there will be no unforeseen contingencies. The previous two subsections suggest that these conditions will be quite restrictive in general games. Some results may be obtained, though, in specific classes of games.

Appendix

Proof of Proposition 3.1. Define the following subset of $\mathcal{A} \times \Delta(\mathcal{A})$:

$$B_1 := \{ \ (A, \mu_0) \in \mathcal{A} \times \Delta(\mathcal{A}) : A' \subseteq A \ ; \ \Gamma(\mu_0) = \{A'\} \ \}, \qquad (R_2) + (KR_1)$$

The second order beliefs in $\Delta(B_1)$ are those that satisfy (R_2) and knowledge of (R_1).

First we show that B_1 is Polish. Since $\mathcal{A}$ is finite – endowed with the discrete topology – it is Hausdorff, implying that every singleton $\{A'\}$ is closed in $\mathcal{A}$, and therefore Polish. Now, B_1 can be rewritten as

$$\bigcup_{A\in\mathcal{A}} \bigcup_{A'\subseteq A} \{ \ (A,\mu_0) \in \mathcal{A}\times\Delta(\mathcal{A}) : \Gamma(\mu_0) = \{A'\} \ \} \ .$$

Since $\{A'\}$ is closed in $\mathcal{A}$ it follows that $\{\mu_0 \in \Delta(\mathcal{A}) : \mu_0(A') \geq 1\}$ is closed in $\Delta(\mathcal{A})$ (Aliprantis and Border, 1994; Corollary 15.6). In addition – as we have already shown – $\{A\}$ is closed in $\mathcal{A}$. Therefore, $\{ \ (A,\mu_0) \in \mathcal{A}\times\Delta(\mathcal{A}) : \Gamma(\mu_0) = \{A'\} \ \}$ is closed in $\mathcal{A}\times\Delta(\mathcal{A})$ and therefore Polish, implying that it is also G_δ. Since the finite union of G_δ sets is also G_δ, it follows that B_1 is Polish.

Now, let $B_k := B_{k-1}\times\Delta(B_{k-1})$ for all $k>1$, and consider belief hierarchies such that μ_0 is such that $\Gamma(\mu_0)$ is a singleton, and also $\mu_k \in \Delta(B_k)$ for $k\geq 1$. These belief hierarchies form the space Θ_2, i.e., they satisfy $(R_1)-(R_2)$ and common knowledge of $(R_1)-(R_2)$.

Since, B_k is Polish for every $k\geq 0$, we can apply Lemma 1 from Brandenburger and Dekel (1993), implying that there is a homeomorphism $\Theta_3^* \mapsto \Delta(\Theta_2)$. Finally, from Proposition 2 in Brandenburger and Dekel (1993) it follows that there is a homeomorphism $g_3 : \Theta_3 \to \Delta(\Theta_3)$, which completes the proof. □

Proof of Proposition 3.2. The proof is similar to the one of Proposition 3.1: Define the following subset of $\mathcal{A}\times\left(\times_{h\in H}\Delta(\mathcal{A})\right)$:

$$\begin{aligned}
\overrightarrow{B_1} :=& \{ \ \left(A,(\overrightarrow{\mu_0}^h;h\in H)\right) \in \mathcal{A}\times\left(\times_{h\in H}\Delta(\mathcal{A})\right) : h=(s(1),...,s(t)) \text{ and } \Gamma(\overrightarrow{\mu_0}^h)=\{A_h\} \Rightarrow \\
&\quad \{s_1(\tau)\}\times\{s_2(\tau)\}\subseteq A_h, \text{ for all } \tau=1,...,t, \text{ and all } h\in H \ \} && (R_2')+(KR_1') \\
&\cap \{ \ \left(A,(\overrightarrow{\mu_0}^h;h\in H)\right) \in \mathcal{A}\times\left(\times_{h\in H}\Delta(\mathcal{A})\right) : A_h\subseteq A, \text{ for all } h\in H \ ; \ \Gamma(\overrightarrow{\mu_0}^h)=\{A_h\} \ \} && (R_3')+(KR_1') \\
&\cap \{ \ \left(A,(\overrightarrow{\mu_0}^h;h\in H)\right) \in \mathcal{A}\times\left(\times_{h\in H}\Delta(\mathcal{A})\right) : A_h\subseteq A_{h'}, \text{ for all } h\in H \text{ and all } h'\in H(h) \ ; \\
&\quad \Gamma(\overrightarrow{\mu_0}^h)=\{A_h\} \text{ and } \Gamma(\overrightarrow{\mu_0}^h)=\{A_h\} \ \} && (R_4')+(KR_1')
\end{aligned}$$

The second order beliefs in $\Delta(\overrightarrow{B_1})$ are those that satisfy $(R_2')-(R_4')$ and knowledge of (R_1').

First we show that $\overrightarrow{B_1}$ is Polish. Since, (R_1') is known, every $\Gamma(\overrightarrow{\mu_0}^h)$ is a singleton, and thus there are finitely many elements in $\overrightarrow{B_1}$. We show that each one of them is a Polish subspace of $\mathcal{A}\times\left(\times_{h\in H}\Delta(\mathcal{A})\right)$, and therefore G_δ, which implies that their union is also G_δ, finally implying that $\overrightarrow{B_1}$ is Polish. Take any $\left(A,(\overrightarrow{\mu_0}^h;h\in H)\right)$ such that $\Gamma(\overrightarrow{\mu_0}^h)=\{A_h\}$ for all $h\in H$. Then, $\{\overrightarrow{\mu_0}^h \in \Delta(\mathcal{A}) : \overrightarrow{\mu_0}^h(A_h)\geq 1\}$ is closed in $\Delta(\mathcal{A})$ for all $h\in H$ (Aliprantis and Border, 1994; Corollary 15.6). In addition, $\{A\}$ is closed in $\mathcal{A}$, and therefore $\left(A,(\overrightarrow{\mu_0}^h;h\in H)\right)$ is closed in $\mathcal{A}\times\left(\times_{h\in H}\Delta(\mathcal{A})\right)$, which proves that $\overrightarrow{B_1}$ is Polish.

Now, let $\overrightarrow{B_k} := \overrightarrow{B_{k-1}}\times\Delta(\overrightarrow{B_{k-1}})$ for all $k>1$, and consider belief hierarchies such that $\overrightarrow{\mu_0}$ is such that $\Gamma(\overrightarrow{\mu_0})$ is a singleton, and also $\overrightarrow{\mu_k}\in\Delta(\overrightarrow{B_k})$ for $k\geq 1$. These belief hierarchies form the space $\overrightarrow{\Theta_4}$, i.e., they satisfy $(R_1')-(R_4')$ and common knowledge of $(R_1')-(R_4')$.

Since, $\overrightarrow{B_k}$ is Polish for every $k\geq 0$, we can apply Lemma 1 from Brandenburger and Dekel (1993), implying that there is a homeomorphism $\overrightarrow{\Theta_5}^* \mapsto \Delta(\times_{h\in H}\overrightarrow{\Theta_4})$. Finally, from Proposition 2 in Brandenburger and Dekel (1993) it follows that there is a homeomorphism $\overrightarrow{g_5} : \overrightarrow{\Theta_5} \to \Delta(\times_{h\in H}\overrightarrow{\Theta_5})$. □

Proof of Proposition 3.3. It follows directly from a standard fixed-point-theorem argument. □

Proof of Proposition 3.4. Observe that i at $\overrightarrow{\theta_i}$ knows j's hierarchy of beliefs at h given that j is of type $\overrightarrow{\theta_j}$. Then i knows what j believes that it is rationalizable. Therefore, i believes that all strategies that can be rationalized under some $\overrightarrow{\theta_j} \in \Gamma(\overrightarrow{g_i}(\overrightarrow{\theta_i}))$ are rationalizable, which completes the proof. □

Proof of Proposition 4.1

(i) Consider a pair of pure actions $(s_1, s_2) \notin a^*$. Since (s_1, s_2) is not distributive knowledge at t^* (i.e. no player is aware of it), either agent can become aware of it if and only if she observes non rationalizable behavior at some $t \geq t^*$. But since $a(\overrightarrow{\theta}_1, h_{t^*}) = a(\overrightarrow{\theta}_2, h_{t^*})$ at t^* every behavior on the equilibrium path is obviously rationalizable. But then at $t^* + 1$ we have again that $a(\overrightarrow{\theta}_1, h_{t^*+1}) = a(\overrightarrow{\theta}_2, h_{t^*+1})$ and hence again every behavior is rationalizable etc..[12]

(ii) Once S^{NE} is first-order mutual knowledge no action pair in $S \backslash S^{NE}$ can be sustained on the equilibrium path. This follows from Benoit and Krishna (1987) together with the following observation. If both players know exactly S^{NE} we are in case (i). Assume thus that one players knows strictly more than S^{NE}. If she only knows more in her own action set (S_i^{NE}), she will never choose such an action. If she knows more also on the opponent's action set, she might want to induce the opponent to learn about a new action. Since she knows, though, that the opponent knows S^{NE} and since any such desired outcome must be in $S \backslash S^{NE}$ she knows that he will not choose such an action. Anticipating this she has no incentive to choose any action outside S_i^{NE}.

(iii) Next we prove necessity using the following game as a counterexample

	X	Y	Z
A	$1, 1$	$0, 0$	$0, 0$
B	$0, 0$	$1, 2$	$\mathbf{100, 100}$
C	$-10, 1$	$-10, 2$	$-10, 3$

.

Note that $S^{NE} = \{A, B\} \times \{X, Y, Z\}$. Assume that $a(\overrightarrow{\theta}_{CP}) = S^{NE}$ (here $\overrightarrow{\theta}_{CP}$ denotes the type of the CP) and $a(\overrightarrow{\theta}_{RP}) = S$ and furthermore that $\text{marg}_{h_t} g(\overrightarrow{\theta} \mid \overrightarrow{\theta}_{RP}) = 1$ where $a(\overrightarrow{\theta}) = \{A, B, C\} \times \{X, Y\}$. (So both players know S^{NE} but this is not first-order mutual knowledge). With these types the RP has an incentive (for ε large enough) to choose C at time t in order for the CP to learn Z. But then the CP will learn C at $t + 1$ hence awareness had not converged at t.

Proof of Proposition 4.2

(i) First we show necessity. Denote by $\overrightarrow{\Theta}_i^{=0}$the subspace of $\overrightarrow{\Theta}_i^0$ s.t. $\forall \overrightarrow{\theta}_i^{=0} \in \overrightarrow{\Theta}_i^{=0} : \overrightarrow{\theta} \in \Gamma\left(\overrightarrow{g}_i(\overrightarrow{\theta}_i^{=0})\right)$ $\Rightarrow a(\overrightarrow{\theta}) \subset S$. Note that $\overrightarrow{\Theta}_i^{=0} \cup \overrightarrow{\Theta}_i^1 = \overrightarrow{\Theta}_i^0$, i.e. we restrict the domain of $\overrightarrow{g}_i$ to exclude first-order mu-

[12]In fact under this condition subjective rationalizability and objective rationalizability in the restricted game with action set a^* do coincide.

tual knowledge. By definition (since $\forall \overrightarrow{\theta} \in \Gamma\left(\overrightarrow{g}_i(\overrightarrow{\theta}_i^{=0})\right) : a(\overrightarrow{\theta}) \subset S$) both players will choose a best response to beliefs contained in $\Delta\left(S'\right)$ for some $S' \subset S$. Since we can choose $\overrightarrow{\theta}_i^{=0} \in \overrightarrow{\Theta}_i^{=0}$ freely it is always possible to find a stage game G and realized types $\widehat{\theta}_i^{=0}$ such that a NE of the game with type space $\left(\overrightarrow{\Theta}_i^{\infty}\right)_{i=1,2}$ cannot be induced on S'.

(ii) Denote by $\overrightarrow{\Theta}_i^{=1}$the subspace of $\overrightarrow{\Theta}_i^{1}$ s.t. $\forall \overrightarrow{\theta}_i^{=1} \in \overrightarrow{\Theta}_i^{=1}, \forall \overrightarrow{\theta}^{\Gamma} \in \Gamma\left(\overrightarrow{g}_i(\overrightarrow{\theta}_i^{=1})\right) : \overrightarrow{\theta} \in \Gamma\left(\overrightarrow{g}_i(\overrightarrow{\theta}_i^{\Gamma})\right) \Rightarrow a(\overrightarrow{\theta}) \subset S$. Recursively we define all $\overrightarrow{\Theta}_i^{=k}$ in this way. For sufficiency we can then simply note that whenever $\overrightarrow{g}_i$ maps into $\Delta\left(\times_{h\in\mathcal{H}}\left(\overrightarrow{\Theta}_i^{=1}\right)\right)$ which it does if the action set is exactly first-order mutual knowledge, then $\forall \overrightarrow{\theta} \in \overrightarrow{\Theta}_i^{=1} : a(\overrightarrow{\theta}) = S_{-i}$ and hence any NE-path can be induced by choosing the appropriate belief on $\Delta(S_{-i})$, which will not be contradicted on the equ. path since it is a NE.

Proof of Proposition 4.3

(i) Again we show first necessity. Consider the type space defined by $\left(\overrightarrow{\Theta}_i^{1}\right)_{i=1,2}$. Then we know (by Proposition 4.2) that all NE-paths of the game with type space $\left(\overrightarrow{\Theta}_i^{\infty}\right)_{i=1,2}$ can be induced. It will be possible, though, to find games s.t. some SPNE-path from the game with type space $\left(\overrightarrow{\Theta}_i^{\infty}\right)_{i=1,2}$ cannot be induced in a (subjective) SPNE for some types in $\left(\overrightarrow{\Theta}_i^{1}\right)_{i=1,2}$. In particular consider types $\overrightarrow{\theta}_i \in \overrightarrow{\Theta}_i^{1}$ s.t. there are two different histories: h_t s.t. $\mathrm{marg}_{h_t}\, \overrightarrow{g}_i(\overrightarrow{\theta}_i) \in \Delta\left(\overrightarrow{\Theta}_{-i}^{=0}\right)$ and h_t' s.t. $\mathrm{marg}_{h_t'}\, \overrightarrow{g}_i(\overrightarrow{\theta}_i) \in \Delta\left(\overrightarrow{\Theta}_{-i}^{=1}\right)$ (as defined in the proof of Proposition 4.2 (ii)). But given this we know (from Proposition 4.2) that we cannot guarantee any NE-path of the common knowledge case to be induced in the "subjective subgame" following history h_t. In particular this can be the case if a Nash action s_i^* is not contained in $\bigcup_{\overrightarrow{\theta}_i^{\Gamma} \in \Gamma\left(\mathrm{marg}_{h_t}\, \overrightarrow{g}_{-i}(\overrightarrow{\theta}_{-i})\right)} a(\overrightarrow{\theta}_i^{\Gamma})$. Of course now such a history h_t can only be reached (by restriction (R5$'$)) if s_i^* is not observed at any time $1, ..., t-1$. This will be the case whenever it is optimal for player i to induce the equilibrium path containing h_t rather than h_t' and (of course whenever player $-i$ has no incentives to deviate). Now by choosing action s_i^* in such a way that the NE-path including s^* is on the continuation path for the SPNE in question and choosing the game payoffs appropriately such a game can always be constructed.

(ii) To establish sufficiency note that if the relevant type space is $\left(\overrightarrow{\Theta}_i^{2}\right)_{i=1,2}$ then the function $\overrightarrow{g}_i$ will map into $\Delta\left(\times_{h\in\mathcal{H}}\left(\overrightarrow{\Theta}_{-i}^{1}\right)\right)$. But this means that in any "subjective subgame" all NE-paths of the game with common knowledge can be recovered (by Proposition 4.2). Consequently all SPNE from the game with type space $\left(\overrightarrow{\Theta}_i^{\infty}\right)_{i=1,2}$ can be recovered.

Proof of Proposition 4.4

(i) Let us first show that not all NE of the game with type space $\left(\overrightarrow{\Theta}_i^{0}, \overrightarrow{g}_i\right)_{i=1,2}$ are (objectively) individually rational. Consider a stage game G with two NE (s_1^*, s_2^*) and $(\widehat{s}_1, \widehat{s}_2)$ s.t. $v_i(s_1^*, s_2^*) - v_i(\widehat{s}_1, \widehat{s}_2) =$

x_i, $\forall i = 1, 2$. Assume $x_1 > 0$ and $x_2 < 0$ and also that s_1^* is not rationalizable in the reduced game with action space $S' = S_1 \times (S_2 \backslash s_2^*)$. Then, player 1 will choose s_1^* at $t = 1$ even if he believes that $\bigcup_{\overrightarrow{\theta}_1^\Gamma \in \Gamma\left(\overrightarrow{g}_1(\overrightarrow{\theta}_1)\right)} a(\overrightarrow{\theta}_1^\Gamma) = S'$. Denote by s_2' the best response of player 2 to s_1^* in the game with action space S' and define $v_1(\widehat{s}_1, \widehat{s}_2) - v_1(s_1^*, s_2') = y_1 < 0$. Then this can be the case e.g. if $\varepsilon\left((T-1)x_1 + y_1\right) + (1-\varepsilon)\left(y_1 + \overline{v}\right) > 0$, where $\overline{v}$ is the continuation payoff in the game starting at $t = 2$ when no discovery occured. Obviously for every ε game payoffs can be found that satisfy this equation. This can be a Nash equilibrium since $x_2 < 0$.

(ii) The fact that not all individually rational payoff vectors can be recovered follows directly from Proposition 4.2.

References

ALIPRANTIS, C., BORDER, K. (1994). *Infinite dimensional analysis.* Springer Verlag, Berlin.

AUMANN, R.J., MASCHLER, M.B. (1995). *Repeated games with incomplete information.* MIT Press, Cambridge.

AUMANN, R.J., BRANDENBURGER, A. (1995). *Epistemic Conditions for Nash Equilibrium.* Econometrica 63, 1161-1180.

BENOIT, J.P. AND V. KRISHNA (1985). Finitely Repeated Games,*Econometrica,* 53(4), 905-922.

BILLINGSLEY, P. (1995). *Probability and measure.* Wiley & Sons, New York.

BRANDENBURGER, A., DEKEL, E. (1987). Common knowledge with probability 1. *Journal of Mathematical Economics* 16, 237–245.

——— (1993). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory* 59, 189–198.

COPIC, J.. AND A. GALEOTTI, (2007). Awareness Equilibrium, mimeo University of Essex.

DEKEL, E., LIPMAN, B., RUSTICHINI, A. (1998). Standard state-space models preclude unawareness. *Econometrica* 66, 159–173.

FAGIN, R., HALPERN, J.Y. (1988). Belief, awareness and limited reasoning. *Artificial Intelligence* 34, 39–76.

FEINBERG, Y.. (2004). Subjective reasoning – Games with unawareness. Discussion paper, Stanford University.

GRANT, S., QUIGGIN, J. (2007). Awareness and discovery. Discussion paper, Rice University.

HALPERN, J. (2001). Alternative semantics for unawareness. *Games and Economic Behavior* 37, 321–339.

HALPERN, J.Y., RÊGO, L.C. (2006). Extensive games with possibly unaware players. *Proceedings Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, 744–751.

HARSANYI, J. (1967-68). Games with incomplete information played by Bayesian players, I-III. *Management Science* 14, 159–182, 320–334, 486–502.

HEIFETZ, A., MEIER, M., SCHIPPER, B. (2007). Unawareness, beliefs and games. GESY Discussion paper 196.

——— (2008). A canonical model of interactive unawareness. *Games and Economic Behavior*, 304-324.

——— (2009). Dynamic unawareness and rationalizable behavior. Discussion paper, UC Davis.

KECHRIS, A. (1995). *Classical descriptive set theory.* Springer Verlag, Berlin.

KREPS, D.M., WILSON, R.B. (1982). Sequential equilibria. *Econometrica* 50, 863–894.

LEE, J. (2008). Unforeseen Contingency and Renegotiation with Asymmetric Information. *Economic Journal 188*, 678-694.

LI, J. (2008). Information structures with unawareness. *Journal of Economic Theory*, forthcoming.

MERTENS, J.F., ZAMIR, S. (1985). Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory* 14, 1–29.

PARTHASARATHY, K.R. (1967). *Probability measures on metric spaces.* AMS Chelsea Publishing, Providence, Rhode Island.

SADZIK, T. (2006). Knowledge, Awareness and Probabilistic Belief, mimeo Graduate School of Business, Stanford University.

SINISCALCHI, M. (2007). Epistemic game theory: beliefs and types. *The New Palgrave Dictionary of Economics*, edited by L. Blume and S. Durlauf, McMillan, forthcoming.