Universidad Nacional de Rosario

Tesina de grado
para la obtención del grado de
Licenciado en Ciencias de la Computación

Vale por un título

Autor:
Tu persona

Directores:
XX
YY

Departamento de Ciencias de la Computación
Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Av. Pellegrini 250, Rosario, Santa Fe, Argentina

9 de abril de 2025

May the force be with you.

YODA (Y PARTICULARMENTE SEBA EN ESTE MOMENTO)

Agradecimientos

Quiero agradecer a quien quiera agradecer. (Por ejemplo, a la universidad pública).

Resumen

Versión cortita de tu tesina.

Índice general

Capítulo 1

Introducción

1.1. Contribuciones

...

1.2. Organización del trabajo

- En el **Capítulo ??**, presentaré ...;
- En el **Capítulo 3**, extenderé y explicaré en detalle los métodos de esta tesina ...;
- En el **Capítulo 4**, presentaré mis resultados y conseguidos ...;
- Finalmente, en el **Capítulo 5**, concluyo con un resumen de los aportes realizados en esta tesina, menciono las implicancias de esta investigación y hallazgos realizados, y sugiero potenciales caminos para futuras investigaciones.

Capítulo 2

Preliminares

2.1. sigma algebras

2.2. politopos

2.3. LTL

Capítulo 3

Verificación de modelos con juegos

3.1. Sobre la verificación de modelos

blabla

problema de sintesis de church

idea de la necesidad de juegos y nuestro entendimiento de ellxs

¿Que preguntas nos hacemos cuando trabajamos con juegos? Mas que nada con juegos estocasticos tenemos las preguntas que se hace Kucera, las que presenta Chatterjee y se puede investigar mas Capaz esto iría mas en la introducción para platear la idea de trabajo

3.2. Cadenas de Markov

Baier and Katoen

3.3. Procesos de Decisión de Markov

Baier and Katoen + Tesis de l de alfaro

3.4. Juegos deterministas

Banerjee ->lo del ejemplito lindo

3.5. Juegos estocásticos

3.5.1. Una breve historia sobre los juegos estocásticos

Enfoque mc ->mdp ->sg ->psg
Enfoque juegos deterministas ->juegos estocasticos (Shapley etc)

3.5.2. Draft

Cosas de Kucera y de la parte preliminar del paper de Pedro.

Me gustaría incluir cosas de Kucera/Banerjee y sus ideas de «1» val etc (concepto de distintas maneras de ganar)

Capaz algo de omega regular stochastic games y qsy

Objetivos

distintos tipos y organizacion / ver de que manera presentarlos

3.6. Juegos Estocásticos Politópicos - PSGs

3.6.1. Draft

Un juego estocástico politópico se caracteriza a través de una estructura que contiene un conjunto finito de estados divididos en dos conjuntos, cada uno perteneciente a un jugador diferente. Además, a cada estado se le asigna un conjunto finito de politopos convexos de distribuciones de probabilidad sobre los estados. La definición formal es como sigue:

Definición 3.6.1 (PSG)**.** *Un juego estocástico politópico (abreviado PSG, por sus siglas en inglés) es una estructura* $\mathcal{K} = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \Theta)$ *tal que* $\mathcal{S}$ *es un conjunto finito de estados particionado en* $\mathcal{S} = \mathcal{S}_\Box \biguplus \mathcal{S}_\Diamond$ *y* $\Theta : \mathcal{S} \to \mathcal{P}_f(\mathrm{DPoly}(\mathcal{S}))$.

La idea de un PSG es la esperada: en un estado $s \in \mathcal{S}_i$ (para $i \in \{\Box, \Diamond\}$), el jugador i elige jugar un politopo $K \in \Theta(s)$ y una distribución $\mu \in K$. El siguiente estado s' se selecciona de acuerdo con la distribución μ, y el juego continúa desde s' repitiendo el mismo proceso.

El desarrollo de un juego estocástico politópico se interpreta en términos de un juego estocástico donde el número de transiciones salientes desde los estados de los jugadores puede ser no numerable. Formalmente, la interpretación de un PSG es la siguiente.

Definición 3.6.2 (Interpretación de un PSG)**.** *La interpretación del juego estocástico politópico* $\mathcal{K}$ *se define por el juego estocástico* $\mathcal{G}_\mathcal{K} = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \mathcal{A}, \theta)$*, donde* $\mathcal{A} = \bigcup_{s \in \mathcal{S}} \Theta(s) \times \mathrm{Dist}(\mathcal{S})$ *y*

$$\theta(s, (K, \mu), s') = \begin{cases} \mu(s') & \textit{si } K \in \Theta(s) \textit{ y } \mu \in K \\ 0 & \textit{en otro caso.} \end{cases}$$

Nótese que el conjunto de acciones $\mathcal{A}$ puede ser no numerable, al igual que cada conjunto $\mathcal{A}(s) = \bigcup_{K \in \Theta(s)} \{K\} \times K$ (pares politopo alcanzable de s y elemento del politopo -distribución sobre el politopo). Por lo tanto, necesitamos extender las estrategias a este dominio no numerable, que debe estar dotado de una σ-álgebra adecuada.

Para esto, utilizamos una construcción estándar para darle a $\mathrm{Dist}(S)$ una σ-álgebra: $\Sigma_{\mathrm{Dist}(S)}$ se define como la σ-álgebra más pequeña que contiene los conjuntos $\{\mu \in \mathrm{Dist}(S) \mid \mu(S) \geq p\}$ para todo $S \subseteq S$ y $p \in [0,1]$. Ahora, dotamos a $\mathcal{A}$ con la σ-álgebra producto $\Sigma_{\mathcal{A}} = \mathscr{P}\left(\bigcup_{s \in S} \Theta(s)\right) \otimes \Sigma_{\mathrm{Dist}(S)}$, y llamamos $\mathrm{PMeas}(A)$ al conjunto de todas las medidas de probabilidad sobre $\Sigma_{\mathcal{A}}$. No es difícil comprobar que cada conjunto de acciones habilitadas $\mathcal{A}(s)$ es medible (es decir, $\mathcal{A}(s) \in \Sigma_{\mathcal{A}}$) y que la función $\theta(s, \cdot, s')$ es medible (es decir, $\{a \in \mathcal{A} \mid \theta(s, a, s') \leq p\} \in \Sigma_{\mathcal{A}}$ para todo $p \in [0,1]$).

Ahora bien, para definir formalmente las estrategias, introduciremos el concepto de comportamiento en PSGs. Esto difiere de la formulación original hecha para PSGs en [1], pero no afecta a los resultados que se mostrarán en este trabajo.

Definición 3.6.3 (Comportamiento en un PSG)**.** *Un comportamiento en un PSG es una secuencia infinita que alterna entre estados y conjuntos resultado. Formalmente un comportamiento es un* $\omega = (s_0, V_0, s_1, V_1, \ldots, s')$ *tal que* $s' \in \mathcal{S}$*,* $s_i \in \mathcal{S}$*,* $V_i \in V_{s_i}$ *y* $s_{i+1} \in V_i$ *para todo* $i \geq 0$*. (Siento que es útil que quede que el último elemento debería pertenecer a S)*

Notaremos al conjunto de todos los comportamientos de un PSG como Ω*. Dado un estado* s*, indicaremos con* Ω_s *el conjunto de los comportamientos que se originan en* s*, y con* Ω_s^n *al conjunto de los comportamientos que se originan en* s *y tiene un total de* n *estados.*

Entonces, ahora sí podemos extender el concepto de estrategia en un PSG:

Definición 3.6.4 (Estrategia en un PSG)**.** *Una estrategia* π_i *para el jugador* i *(*$i \in \{\Box, \Diamond\}$*) en un PSG será una función* $\pi_i : \Omega\mathcal{S}_i \to \mathrm{PMeas}(\mathcal{A})$ *que asigna una medida de probabilidad a cada comportamiento* ωs *tal que* $\pi_i(\omega s)(\mathcal{A}(s)) = 1$*.*

En realidad habría que ir al concepto de acciones. Modificar, pero igual después preguntar bien por qué no va este enfoque de conjuntos soporte -además de porque es raro-. Habría que agregar alguna justificación de esta idea de comportamientos, ¿no?

Medida de probabilidad que no va a ir probablemente

Con la formalización del concepto de estrategia ahora podemos presentar formalmente la definición de $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}$, la medida de probabilidad definida por la cadena de Markov dada por $\mathcal{G}_\mathcal{K}$ y las estrategias $\pi_\Box$ y $\pi_\Diamond$.
Para eso, primero para cada $n \geq 0$ y $s \in \mathcal{S}$ definimos $\mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s} : \Omega^{n+1} \to [0,1]$ para todo $s' \in \mathcal{S}$ y $\hat{\omega} \in \Omega^{n+1}_s$ inductivamente como sigue:

$$\mathbb{P}^{\pi_\Box,\pi_\Diamond,0}_{\mathcal{G}_\mathcal{K},s}(s') = \delta_s(s')$$

$$\mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat{\omega}s_n V' s') = \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\omega}s_n) \int_{\{a \in A(s_n) | \mathrm{sop}(a) = V'\}} \theta(s_n, a, s')\, d(\pi_i(\hat{\omega}s_n)(a))$$
$$\text{si } s_n \in \mathcal{S}_i \text{ con } i \in \{\Box, \Diamond\}$$

(sop(a) = V' significa que el soporte de a es V')
(capaz restringirse sobre las acciones con soporte V'? No sé bien cómo va a quedar esto)
y extendemos $\mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s} : \mathcal{P}(\Omega^{n+1}) \to [0,1]$ a conjuntos como la suma de la medida sobre los elementos del conjunto.
Sea Σ_Ω la σ-álgebra discreta sobre Ω (pues tanto $\mathcal{S}$ como los conjuntos soportes serán conjuntos finitos) y Σ_{Ω^ω} la σ-álgebra producto usual sobre Ω^ω. Por el teorema de extensión de Carathéodory, $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s} : \Sigma_{\Omega^\omega} \to [0,1]$ se define como la única medida de probabilidad tal que para todo $n \geq 0$, y $SV_i \in \Sigma_\Omega$, $0 \leq i \leq n$,

$$\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(SV_0 \times \cdots \times SV_n \times \Omega^\omega) = \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(SV_0 \times \cdots \times SV_n).$$

3.7. Procesos de Decisión de Markov Politópicos - PMDPs

aclarar que son aporte las cosas que se ven aca

3.7.1. Draft

Definición 3.7.1. *Un proceso de decisión de Markov politópico (PMDP por sus siglas en inglés) es una tupla $\mathcal{M} = (\mathcal{S}, Act, \theta')$ donde S es un conjunto finito de estados, Act es un conjunto de pares (politopo, distribución) y $\theta' : \mathcal{S} \times Act \times \mathcal{S} \to [0,1]$ es la función de transición entre estados. También podemos definir a un PDMP como un PSG donde $\mathcal{S}_\Box = \emptyset$ o $\mathcal{S}_\Diamond = \emptyset$.*

Antes que nada, presentamos el concepto de conjunto resultado:

Definición 3.7.2 (conjunto resultado)**.** *Para cada $s \in \mathcal{S}$ notaremos con V_s al conjunto de todos los subconjuntos de posibles estados siguientes que permiten las distintas (e* *infinitas) distribuciones desde s. Es decir,*

$$V_s = \{V' \subseteq \mathcal{S} | \exists \mu \in K, K \in \Theta(s) \text{ tal que } \mu(V') = 1 \text{ y } \forall s' \in V', \mu(s') > 0\} \tag{3.1}$$

Nótese que como $\mathcal{S}$ es finito, V_s también lo será, así como cada V'.
*Llamaremos **conjunto resultado** a cada elemento $V' \in V_s$.*

También introducimos el concepto de *comportamiento* en un PMDP. Un comportamiento será una secuencia alternante de estados y conjuntos de próximos estados posibles, que refleja un proceso de selección de dos pasos. Desde un estado s, primero, el jugador selecciona un politopo y una distribución cuyo conjunto soporte será un $V \in V_s$, y luego el próximo estado s' será elegido probabilísticamente en base a la distribución seleccionada. Presentamos la definición formal de la siguiente manera:

Si cambiamos lo de PSGs y comportamientos hay que cambiar esta subsección completa también.

Definición 3.7.3 (comportamiento)**.** *Un comportamiento en un PMDP es una secuencia infinita $\omega = (s_0, V_0, s_1, V_1, ...)$ tal que $s_i \in \mathcal{S}$, $V_i \in V_{s_i}$ y $s_{i+1} \in V_i$ para todo $i \geq 0$.*

Dado un estado s, indicamos con Ω_s el conjunto de los comportamientos que se originan en s.

Las estrategias en los PMDPs se definen de manera análoga a las estrategias en juegos estocásticos politópicos, mientras que las definiciones de conjuntos estado-acción, sub-MDPs y componentes finales de los procesos de decisión de Markov tradicionales se extienden a los PMDPs de la siguiente manera:

Definición 3.7.4 (conjuntos estado-resultado y sub-PMDPs)**.** *Dado un PMDP $\mathcal{M} = (S, Act, \theta')$ un conjunto estado-resultado es un subconjunto $\chi \subseteq \{(s, V') \mid s \in S \wedge V' \in V_s\}$. Un sub-PMDP es un par (C, D), donde $C \subseteq S$ y D es una función que asocia a cada $s \in C$ un conjunto $D(s) \subseteq V_s$ de subconjuntos de estados próximos posibles. Hay una relación uno-a-uno entre sub-PMDPs y conjuntos de estado-acción:*

- *dado un conjunto estado-resultado χ, denotamos $sub(\chi) = (C, D)$ al sub-PMDP definido por:*

$$C = \{s \mid \exists V'.(s, V') \in \chi\} \quad D(s) = \{V' \mid (s, V') \in \chi\}$$

- *dado un sub-PMDP (C, D), denotamos por $er(C, D) = \{(s, V') \mid s \in C \wedge V' \in D(s)\}$ al conjunto estado-resultado correspondiente a (C, D).*

Si definimos para cada V', un vértice único nuevo $v_{V'}$ podemos ver que cada sub-PMDP (C, D) induce una *relación de aristas*: hay una arista $(s, v_{V'})$ de $s \in C$ a $v_{V'}$ para cada $V' \in D(s)$ y hay una arista de $(v_{V'}, t)$ de $v_{V'}$ con $V' \in D(s)$ a $t \in \mathcal{S}$ sii es posible ir de s a t en un paso con probabilidad positiva utilizando una acción cuyo conjunto resultado sea V'. La definición formal es como sigue:

Definición 3.7.5 (relación de aristas ρ)**.** *Para un sub-PMDP, definimos la relación $\rho_{(C,D)}$ como*

$$\rho_{(C,D)} = \{(s, v_{V'}) \mid \exists V' \in D(s)\} \cup \{(v_{V'}, t) \mid \exists s \in C. V' \in D(s) \wedge t \in V'\}$$

Definición 3.7.6 (componente final)**.** *Un sub-PMDP es una componente final si:*

- *$V' \subseteq C$ para todo V' tal que existe un $s \in C$ donde $V' \in D(s)$*
- *el grafo $(C, \rho_{(C,D)})$ es fuertemente conexo.*

Llamaremos $EC(\mathcal{M})$ al conjunto de todas las componentes finales en un PMDP $\mathcal{M}$.

Intuitivamente, una componente final representa un conjunto de pares estado-resultado que, una vez en ellos, es posible quedase allí si la estrategia escoge las acciones de manera apropiada. Esta intuición se hará precisa con los siguientes teoremas.

Antes de enunciar estos teoremas, introducimos una abreviatura para el conjunto de estados-resultados que ocurren infinitamente a menudo en un comportamiento dado y una notación para el conjunto (infinito) de acciones con el mismo conjunto resultado.

Definición 3.7.7 (*inf*)**.** *Dado un comportamiento $\omega = (s_0, V'_0, s_1, V'_1, \dots)$ indicamos por*

$$\textit{inf}(\omega) = \{(s, V') \mid s_k = s \wedge V'_k = V' \text{ para infinitos } k \in \mathbb{N}_0\}$$

al conjunto de pares estado-resultado que ocurren infinitas veces en él.

Definición 3.7.8 (A(s, V'))**.** *Notaremos al conjunto de acciones desde un estado $s \in S$ con el mismo conjunto resultado $V' \subseteq V_s$ como:*

$$A(s, V') = \{(K, \mu) \mid K \in \Theta(s) \wedge \mu \in K \wedge \mu(V') = 1 \wedge \forall v' \in V',\ \mu(v') > 0\}$$

Teorema 3.7.1 (estabilidad de componentes finales)**.** *Sea (C, D) una componente final. Entonces, para cada estrategia π existe una estrategia π', que difiere de π solo en C, tal que:*

$$\mathbb{P}^{\pi}_s(\Diamond C) = \mathbb{P}^{\pi'}_s(\Diamond C) = \mathbb{P}^{\pi'}_s(inft(\omega) = er(C, D)) \tag{3.2}$$

para todo $s \in S$.

Demostración. Considérese una estrategia π' definida como sigue para cada secuencia $s_0 \dots s_n$ con $n \geq 0$:

- Si $s_n \in C$, la estrategia asignará probabilidad positiva a una única acción $(K, \mu) \in A(s_n, V')$ para cada conjunto resultado V' (notaremos a esta acción particular con $(K, \mu)_{V'}$), y la probabilidad de elegir cada una de esas acciones se distribuirá de manera uniforme. Es decir,

$$\pi'(s_0 \dots s_n)(K, \mu) = \begin{cases} \frac{1}{|D(s_n)|} & \text{si } (K, \mu) = (K, \mu)_{V'} \text{ para algún } V' \in D(s); \\ 0 & \text{en otro caso} \end{cases}$$

- Si $s_n \notin C$, la estrategia π' coincide con π, i.e.

$$\pi(s_0 \dots s_n)(K, \mu) = \pi'(s_0 \dots s_n)(K, \mu) \quad \forall (K, \mu) \in Act$$

La primera igualdad en 3.2 es una consecuencia del hecho de que π y π' coinciden fuera de C.

Para la segunda igualdad, basta con ver que bajo la estrategia π' una vez que un comportamiento entra a C, nunca sale de C ni se elige una acción que no esté en D. Es más, una vez en C un comportamiento visitará todos los estados de C infinitamente a menudo con probabilidad 1. $\square$

El próximo resultado establece que, para cualquier estado inicial y cualquier estrategia (de memoria finita), un comportamiento terminará con probabilidad 1 en una componente final. Esta es la razón detrás del nombre "componente final".

Teorema 3.7.2 (teorema fundamental de las componentes finales)**.** *Sea $\mathcal{M}$ un PMDP. Para todo $s \in S$, toda estrategia π de memoria finita,*

$$\mathbb{P}^{\pi}_{\mathcal{M},s}(\{\omega \in \mathit{Paths}(s) \mid \mathit{sub}(\mathit{inft}(\omega)) \textit{ es una componente final}\}) = 1$$

Demostración. Consideremos un sub-PMDP (C, D) que no sea una componente final y sea $\Omega_s^{(C,D)} = \{\hat{\omega} \in \Omega_s \mid \mathit{inft}(\hat{\omega}) = \mathit{er}(C, D)\}$ el conjunto de comportamientos cuyo conjunto de pares estado-resultado que se repiten infinitas veces en él forman el sub-PMDP (C, D).

Si podemos mostrar que

$$\mathbb{P}^{\pi}_{\mathcal{M},s}(\{\omega \in \mathit{Paths}(s) \mid \omega \in \Omega_s^{(C,D)}\}) = 0 \tag{3.3}$$

como (C, D) es un sub-PMDP cualquiera y como hay una cantidad finita de sub-PMDPs en $\mathcal{M}$, esto es lo mismo que mostrar que

$$\mathbb{P}^{\pi}_{\mathcal{M},s}(\{\omega \in \mathit{Paths}(s) \mid \mathit{sub}(\mathit{inft}(\omega)) \text{ es una componente final}\}) = 1$$

.

Veamos que vale 3.3, dividiendo en casos según cuál es la condición de la definición 3.7.6 que no se cumple para (C, D):

- Primero, asumamos que existe un $(t, V') \in \mathit{er}(C, D)$ tal que $V' \nsubseteq C$.

  Sabemos que cada comportamiento en $\Omega_s^{(C,D)}$ toma el par estado-resultado (t, V') infinitas veces. Llamemos I al conjunto de índices infinito que representa los momentos en los que se visita (t, V'). Indiquemos con μ_i a la distribución elegida en el momento $i \in I$ y definamos como $r_i = \sum_{u \in C} \mu_i(u)$ a la probabilidad de quedarnos en C en el momento $i \in I$ (que en cada caso será menor a 1 porque $V' \nsubseteq C$).

  Como π es de memoria finita sucederá que π solo puede elegir una cantidad finita de acciones (y, por lo tanto, distribuciones) distintas desde t. Esto hace que el conjunto $R = \{r_i \mid i \in I\}$ tenga un máximo, llamémoslo r.

Para que valga que ω esté en $\Omega_s^{(C,D)}$ tiene que valer que en infinitos momentos i nos quedemos en C. Entonces que vale que $\mathbb{P}_s^\pi(\omega \in \Omega_s^{(C,D)}) < r^k$ para todo $k > 0$ natural. Como sabemos que $r < 1$, tenemos que $\mathbb{P}^\pi_{\mathcal{M},s}(\{\omega \in \mathit{Paths}(s) \mid \omega \in \Omega_s^{(C,D)}\}) = 0$.

- Si no, asumamos que existen $t_1, t_2 \in C$ tales que no hay camino de t_1 a t_2 en $(C, \rho(C,D))$.

  La falta de camino de t_1 a t_2 en $(C, \rho_{(C,D)})$ implica que para cada subsecuencia $s_m V_m s_{m+1} ... s_n$ de comportamiento en $\Omega_s^{(C,D)}$ que vaya de $s_m = t_1$ a $s_n = t_2$, hay 2 opciones:

  1. existe un $j \in [m+1, n-1]$ tal que $s_j \notin C$. Como cada comportamiento en $\Omega_s^{(C,D)}$ contiene infinitas subsecuencias de t_1 a t_2 y tenemos una cantidad finita de estados, sabemos que habrá una cantidad infinita de s_j iguales. Pero si infinitas veces se toma un estado s_j entonces, $s_j \in C$, lo que contradice la hipótesis anterior. Absurdo.
  2. existe un $j \in [m, n-1]$ tal que $V_j \notin D(j)$. Como cada comportamiento en $\Omega_s^{(C,D)}$ contiene infinitas subsecuencias de t_1 a t_2 y tenemos una cantidad finita de conjuntos resultado, sabemos que habrá una cantidad infinita de V_j iguales, con lo que $V_j \in D(j)$, lo que contradice la hipótesis anterior. Absurdo

  Con lo que arribamos a que $\mathbb{P}^\pi_{\mathcal{M},s}(\{\omega \in \mathit{Paths}(s) \mid \omega \in \Omega_s^{(C,D)}\}) = 0$

□

Esto nos permite definir el siguiente corolario útil para el análisis de condiciones de Rabin:

Corolario 3.7.2.1. *Sea $\mathcal{M}$ un PMDP, s un estado en él y sea π una estrategia (lo hacemos con estrategias o planificadores?) en él. Una condición de Rabin $\widetilde{R} = \{\langle G_1, R_1\rangle, ..., \langle G_k, R_k\rangle\}$ se satisface desde s, siguiendo la estrategia π con probabilidad 1 si y solo si para cada componente final U alcanzable desde s, existe un $j \in \{1, 2, ..., k\}$ tal que $U \cap R_j = \emptyset$ y $U \cap G_j \neq \emptyset$.*

Ahora bien, una pregunta muy natural que surge es "¿por qué nos restringimos a estrategias de memoria finita en este último teorema?"

Nota sobre la limitación a estrategias finitas del Teorema 3.7.2

La prueba del teorema 3.7.2 está inspirada en las pruebas realizadas para MDPs (véase [2] [3]). En ellas, la prueba consiste en proponer un sub-PMDP (C, D) arbitrario que no cumple la definición 3.7.6, dividir el análisis por casos de cómo no se cumple la definición de componente final y mostrar que en cada caso la probabilidad de que un comportamiento que siga la estrategia del enunciado genere un sub-PMDP como el propuesto es 0.

Si nos remitimos a la prueba que mostramos en 3.7.2, vemos que el método de prueba se ajusta bien a nuestro caso hasta llegar al primer ítem donde el argumento está explícitamente respaldado en el hecho de que π es de memoria finita. De esta manera, se puede decir que el conjunto de las distintas probabilidades de quedarse en C en los momentos i tiene un máximo, r, lo que permite acotar la productoria $\Pi_{i \in I}\, r_i$ por $\lim_{k\to\infty} r^k$, que sabemos que converge a 0, puesto que $r < 1$.

Si no nos restringimos a estrategias de memoria finita no podemos garantizar esto. Sabemos que la productoria va a estar acotada inferiormente por 0 y superiormente por 1, pero bien podría no converger a 0, sino a algún otro número. Por ejemplo, veamos lo siguiente:

Consideremos un juego estocástico politópico donde tenemos un vértice t y un vértice s. Desde t, existen acciones μ de la forma $\mu(s) = \frac{1}{k^2}$, $\mu(t) = 1 - \frac{1}{k^2}$, por lo que se puede definir una estrategia de memoria infinita π a partir de la cantidad de t que se encuentran en el comportamiento de la siguiente manera:

$$\pi(\omega t)(s) = \frac{1}{(\#_t(\omega))^2}$$
$$\pi(\omega t)(t) = 1 - \frac{1}{(\#_t(\omega))^2}$$

donde $\#_t$ se define inductivamente sobre comportamientos de la siguiente manera:

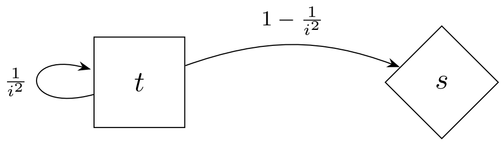

Figura 3.1: Interpretación del juego estocástico con la estrategia π fijada.

$$\#_t(\emptyset) = 0$$
$$\#_t(\omega V' t) = \#_t(\omega) + 1$$
$$\#_t(\omega V' s) = \#_t(\omega)$$

donde ω es un comportamiento, $\emptyset$ representa el comportamiento vacío (no sé qué tan bien estaría esto) y $s \neq t$.

Si, remitiéndonos otra vez a la estrategia de prueba, suponemos que s no forma parte de C, tenemos que la probabilidad de quedarnos en C sería $\Pi_{i\geq 0} 1 - (\frac{1}{i^2})$, que sabemos que converge a $\frac{1}{2}$.

No me gustan mucho estos siguientes párrafos, pero no sé muy bien cómo decir lo que quiero, y siento que vos habías dicho algo más interesante sobre esto.

Ahora bien, es importante aclarar que esto no constituye de ninguna manera un contraejemplo, además de que faltaría formalidad en su presentación, resulta que plantear $s \notin C$ siendo que tengo probabilidad positiva de visitarlo cada vez no tiene mucho sentido. (Esto no sé que tanto es así).

En cualquier caso, podría ser una muestra de que la aplicación de este fórmula de prueba no es útil *directamente* para probar el teorema para estrategias de memoria infinita. Pero resaltamos el directamente porque la productoria constituye una cota a la probabilidad de que $\omega \in \Omega_s^{(C,D)}$, no es directamente el valor de la probabilidad.

(Faltaría agregar figura)

Esto de acá abajo capaz se va

A su vez, el teorema anterior nos permite empezar a considerar especialmente un tipo de propiedades, aquellas que solo dependen del comportamiento asintótico.

Definición 3.7.9 (propiedad límite)**.** *Una propiedad de tiempo lineal se llama propiedad límite de tiempo lineal si para todos los comportamientos* ω, ω' $\omega \in P \wedge \mathit{inf}(\omega) = \mathit{inf}(\omega') \implies \omega' \in P$.

Si T *es un subconjunto de estados para un PMDP* $\mathcal{M}$*, diremos que* $T \models P$ *para alguna propiedad límite* P *sii para todo comportamiento tal que* $\mathit{sub}(\mathit{inf}(\omega)) = (T, A)$ *para algún* A*, vale que* $\omega \in P$.

Por el teorema 3.7.2, solo los conjuntos T donde T es un conjunto de estados de una componente final son relevantes a la hora de analizar las probabilidades de una propiedad límite P. Denotemos U_P a la unión de todos los conjuntos T de todas las componentes finales (T, A) de $\mathcal{M}$ tales que $T \models P$, y V_P la unión de los conjuntos T de todos los $(T, A) \in EC(\mathcal{M})$ tales que $\neg(T \models P)$. Esto nos permite presentar el siguiente interesante resultado:

Teorema 3.7.3. *Sea* $\mathcal{M}$ *un PMDP y sea* P *una propiedad de límite. Entonces existe una estrategia de memoria finita* π *tal que para todo estado* s *de* $\mathcal{M}$*:*

(a) $\sup_\pi \mathbb{P}^\pi_{\mathcal{M},s}(P) = \sup_\pi \mathbb{P}^\pi_{\mathcal{M},s}(\Diamond U_P)$

(b) $\inf_\pi \mathbb{P}^\pi_{\mathcal{M},s}(P) = 1 - \sup_\pi \mathbb{P}^\pi_{\mathcal{M},s}(\Diamond V_P)$

Demostración. Primero, probemos el item *(a)*.

Para cada estrategia tenemos que $\mathbb{P}^\pi_{\mathcal{M},s}(P) \leq \mathbb{P}^\pi_{\mathcal{M},s}(\Diamond U_P)$, puesto que por teorema 3.7.2 vale que:

$$\mathbb{P}^\pi_{\mathcal{M},s}(P) = \mathbb{P}^\pi_{\mathcal{M},s}\{\omega \in \mathit{Paths}(s) \mid \mathit{inf}(\omega) \in EC(\mathcal{M}) \wedge \mathit{inf}(\omega) \models P\}$$

y por definición de U_P, $\omega \models \Diamond U_P$ para cada camino ω tal que $\mathit{inf}(\omega)$ es una componente final y $\mathrm{inf}(\omega) \models P$.

Si podemos probar que existe una estrategia π tal que $\mathbb{P}^\pi_{\mathcal{M},s}(P) = \sup_\pi \mathbb{P}^\pi_{\mathcal{M},s}(\Diamond U_P)$, podemos ver que vale la igualdad, por lo expuesto anteriormente.

Consideremos, entonces, una estrategia sin memoria π_0 que maximiza las probabilidades de llegar a U_P para cada s en $\mathcal{M}$ (esta estrategia existe por resultado X de literatura y el teorema 1 de [1]). Además, para cada componente final (C, D), definamos $\pi_{(C,D)}$: una estrategia que asegura quedarse siempre en C mientras visita infinitamente a menudo todos los estados $s \in C$ (esta estrategia existe por el teorema 3.2). Y, por último, para cada estado $u \in U_P$ seleccionamos una componente final $EC(u) = (C, D)$

tal que $u \in C$ y $C \models P$. Todo esto nos servirá para definir la estrategia π que nos asegura que $\mathbb{P}^{\pi}_{\mathcal{M},s}(P) = \sup_{\pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(\Diamond U_P)$.

Definimos π como la estrategia que primero se comporta como π_0 hasta que llega a un estado $u \in U_P$. Desde ese momento, π se comporta como $\pi_{EC(u)}$. En consecuencia, para esta estrategia, los caminos que eventualmente entran a U_P visitarán todos los estados de una componente final (C, D) tal que $C \models P$ infinitamente a menudo con probabilidad 1. En particular, *inf*$(\pi) \models P$ vale para todo ω que sigue la estrategia π y eventualmente alcanza U_P. Esto lleva a que:

$$\begin{aligned}
\mathbb{P}^{\pi}_{\mathcal{M},s}(P) &= \sum_{u \in U_P} \mathbb{P}^{\pi_0}_{\mathcal{M},s}((\neg U_P)\ \mathcal{U}\ u) \cdot \mathbb{P}^{\pi_{EC(u)}}_{\mathcal{M},u}(P) \\
&= \sum_{u \in U_P} \mathbb{P}^{\pi_0}_{\mathcal{M},s}((\neg U_P)\ \mathcal{U}\ u) \cdot 1 \\
&= \sum_{u \in U_P} \mathbb{P}^{\pi_0}_{\mathcal{M},s}((\neg U_P)\ \mathcal{U}\ u) \\
&= \sup_{\pi'} \mathbb{P}^{\pi'}_{\mathcal{M},s}(\Diamond U_P)
\end{aligned}$$

(Revisar la manera en la que están escritas estas igualdades).

Hemos probado así el ítem *(a)*.

Ahora, probemos el ítem *(b)*. Este deriva del ítem anterior usando el hecho de que las propiedades límite son cerradas bajo negación (es decir, si P es una propiedad límite entonces $\overline{P}$ también lo es) y que $\mathbb{P}^{\pi}_{\mathcal{M},s}(P) = 1 - \mathbb{P}^{\pi}_{\mathcal{M},s}(\overline{P})$ para toda estrategia π. Por lo tanto:

$$\inf_{\pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(P) = 1 - \inf_{\pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(\overline{P}) \tag{3.4}$$

y cualquier estrategia de memoria finita que maximiza la probabilidades para $\overline{P}$, minimiza las probabilidades para P. Para un subconjunto de estados T, tenemos que $T \models \overline{P}$ sii $\neg(T \models P)$. Entonces, el conjunto V_P de todos los estados t que están contenidos en una componente final (T, A) con $\neg(T \models P)$ coincide con el conjunto $U_{\overline{P}}$ que surge de la unión de todos las componentes finales (T, A) donde $T \models \overline{P}$. Y con eso, el ítem *(a)* aplicado a $\overline{P}$ hace que valga:

$$\sup_{\pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(\overline{P}) = \sup_{\pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(\Diamond U_{\overline{P}}) = \sup_{\pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(\Diamond V_P)$$

Y estas igualdades aplicadas a 3,4 nos dan lo que queríamos probar.

□

3.8. Comparación de los distintos juegos

Incluir cuadros comparativos entre juegos // ejemplos para todas las nociones nuevas explicadas

(En general, estaría bueno ver problemas y agregar un apéndice con ejemplos)

Capítulo 4

Respondiendo a las preguntas

4.1. Juegos justos: la rta a la pregunta cualitativa

4.1.1. Draft

Agg cambiando lo de las acciones

4.2. Transformar Rabin a alcanzabilidad

4.2.1. Draft

La idea es intentar hacer algo al estilo

Teorema 4.2.1 (Reducción de Rabin)**.** *Sea $\mathcal{G}_{\mathcal{K}}$ un juego estocástico politópico y sea $\mathcal{H}_{\mathcal{K}}$ su interpretación extrema tal como se presenta en [1]. Si tenemos una propiedad de Rabin R, y su respectivo conjunto de estados almost sure winning W_R entonces valen las siguientes ecuaciones:*

$$
\begin{aligned}
&\inf_{\pi_\Diamond \in \Pi_{\mathcal{G}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(R) = \\
&= \inf_{\pi_\Diamond \in \Pi_{\mathcal{G}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond W_R) = \\
&= \inf_{\pi_\Diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond W_R) = \\
&= \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \inf_{\pi_\Diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Diamond}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond W_R) = \\
&= \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\Diamond \in \Pi_{\mathcal{G}_\mathcal{K},\Diamond}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond W_R) = \\
&= \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\Diamond \in \Pi_{\mathcal{G}_\mathcal{K},\Diamond}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(R)
\end{aligned}
$$

Demostración.

$$
\begin{aligned}
&\inf_{\pi_\Diamond \in \Pi_{\mathcal{G}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(R) \\
&\quad \leq \inf_{\pi_\Diamond \in \Pi_{\mathcal{G}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond W_R) && \text{(por lema 4,2,2)} \\
&\quad \leq \inf_{\pi_\Diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond W_R) && \text{(por Teorema 1)} \\
&\quad \leq \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \inf_{\pi_\Diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Diamond}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond W_R) && \text{(por Teorema 1)} \\
&\quad \leq \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\Diamond \in \Pi_{\mathcal{G}_\mathcal{K},\Diamond}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond W_R) && \text{(por Teorema 1)} \\
&\quad \leq \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\Diamond \in \Pi_{\mathcal{G}_\mathcal{K},\Diamond}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(R) && \text{(por lema 4,2,4)} \\
&\quad \leq \inf_{\pi_\Diamond \in \Pi_{\mathcal{G}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(R) && \text{(por prop de sup e inf)}
\end{aligned}
$$

□

Lema 4.2.2. *Sea $\mathcal{G}_\mathcal{K}$ la interpretación (?) de un juego estocástico politópico y sea R una propiedad límite y W_R su correspondiente conjunto casi seguramente ganador. Entonces vale que*

$$\inf_{\pi_\Diamond \in \Pi_\Diamond} \sup_{\pi_\Box \in \Pi_\Box} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(R) \leq \inf_{\pi_\Diamond \in \Pi_\Diamond} \sup_{\pi_\Box \in \Pi_\Box} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond W_R)$$

Demostración. Supongamos una estrategia $\pi_\Diamond$ arbitraria de memoria finita. Sabemos que $\mathcal{G}_\mathcal{K}$ con una estrategia de memoria finita fijada es un PMDP. Entonces, por lema 4.2.3 vale que para cada estrategia de memoria finita $\pi_\Diamond$

$$\sup_{\pi_\Box \in \Pi_\Box} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(R) = \sup_{\pi_\Box \in \Pi_\Box} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond W_R)$$

lo que implica que (ahora no sé si me convence, tengo que redactar mejor esta justificación) capaz tengo que agg lo de contencion de caminos en P y WP

$$\inf_{\pi_\Diamond \in \Pi_\Diamond} \sup_{\pi_\Box \in \Pi_\Box} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(R) \leq \inf_{\pi_\Diamond \in \Pi_\Diamond} \sup_{\pi_\Box \in \Pi_\Box} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond W_R)$$

□

Lema 4.2.3. *Sea $\mathcal{M}$ un PMDP, π una estrategia en $\mathcal{M}$, s un estado en $\mathcal{M}$, R una propiedad de Rabin y W_R el conjunto casi seguramente ganador de R (en el juego estocástico debería ser, no?), entonces vale que:*

$$\sup_{\pi\in\Pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(R) = \sup_{\pi\in\Pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(\Diamond W_R)$$

Demostración. Por definición de conjunto casi seguramente ganador

$$\mathbb{P}^{\pi}_{\mathcal{M},s}(R) = \mathbb{P}_{\mathcal{M},s}(\{\omega \in \mathit{Paths}(s) \mid \mathit{inf}(\omega) \models R\})$$

Claramente los caminos que hacen que valga la propiedad también son caminos desde donde llego a algún estado desde donde existen estrategias para ganar con probabilidad 1. Por lo tanto, $\mathbb{P}^{\pi}_{\mathcal{M},s}(R) \leq \mathbb{P}^{\pi}_{\mathcal{M},s}(\Diamond W_R)$.

Para ver que vale la igualdad, veremos que existe una estrategia de memoria finita π que hace que $\mathbb{P}^{\pi}_{\mathcal{M},s}(R) = \sup_{\pi\in\Pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(\Diamond W_R)$. Para ello, consideremos la estrategia sin memoria π_0 que maximiza las probabilidades de alcanzar W_R desde todos los estados $s \in \mathcal{M}$ (sabemos que esta estrategia existe por [1, 4]). A su vez, sabemos que para cada estado $s \in W_R$ existe una estrategia π_s que asegura ganar con probabilidad 1 frente al objetivo R.

Sea entonces π la estrategia que primero se comporta como π_0, hasta llegar a un estado t en W_R, y a partir de allí π se comporta como π_t. Con eso tenemos que:

$$\begin{aligned}\mathbb{P}^{\pi}_{\mathcal{M},s}(R) &= \sum_{t\in W_R} \mathbb{P}^{\pi_0}_{\mathcal{M},s}((\neg W_R \,\mathcal{U}\, t)) \cdot \underbrace{\mathbb{P}^{\pi_t}_{\mathcal{M},t}(R)}_{=1} \\ &= \sup_{\pi\in\Pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(W_R)\end{aligned}$$

Como $\sup_{\pi\in\Pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(W_R)$ es una cota superior las probabilidades para R bajo todas las estrategias, con esto podemos concluir la igualdad $\sup_{\pi\in\Pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(R) = \sup_{\pi\in\Pi} \mathbb{P}^{\pi}_{\mathcal{M},s}(\Diamond W_R)$. □

Esto es lo importante que habría que probar y no está probado.

Lema 4.2.4. *Sea $\mathcal{G}_\mathcal{K}$ la interpretación (?) de un juego estocástico politópico y sea R una propiedad límite y W_R su correspondiente conjunto casi seguramente ganador. Entonces vale que*

$$\sup_{\pi_\Box \in \Pi_\Box} \inf_{\pi_\Diamond \in \Pi_\Diamond} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond W_R) \leq \sup_{\pi_\Box \in \Pi_\Box} \inf_{\pi_\Diamond \in \Pi_\Diamond} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(R)$$

Capítulo 5

Conclusiones

5.1. Trabajo Futuro

Referencias

[1] P. F. Castro y P. R. D'Argenio. «Polytopal Stochastic Games». En: (2024). Enviado para publicación.

[2] C. Baier y J.-P. Katoen. *Principles of Model Checking*. Vol. 26202649. The MIT Press, 2008. ISBN: 978-0-262-02649-9.

[3] L. de Alfaro. *Formal Verification of Probabilistic Systems*. Inf. téc. Stanford, CA, USA, 1998.

[4] A. Condon. «The complexity of stochastic games». En: *Information and Computation* 96.2 (1992), págs. 203-224. ISSN: 0890-5401. DOI: `https://doi.org/10.1016/0890-5401(92)90048-K`. URL: `https://www.sciencedirect.com/science/article/pii/089054019290048K`.

Apéndice A

Titulo del Apendice

A.1. Titulo de la seccion