UNIVERSIDAD DE BUENOS AIRES

Facultad de Ciencias Exactas y Naturales

Departamento de Matemática

Tesis de Licenciatura

Teoría de juegos e investigación operativa: Una aplicación en el Aeropuerto de Ezeiza

Emanuel Javier Ferreyra

Director: Guillermo Durán **Codirector: Fernando Ordóñez**

Fecha de Presentación : Marzo 2016

Agradecimientos

A Willy, por la oportunidad de usar la matemática fuera de la facultad, con mucho afecto. Junto con Fernando me ayudaron a abrir caminos y me dejaron enseñanzas que trascienden este trabajo.

A Pinasco, que supo aconsejarme y responder mis consultas en muchas ocasiones, y me introdujo en este tema tan lindo. A él y a Flavia por corregir esta tesis.

A todos los docentes que tuve en la carrera, en particular a Rela y Victoria; a Acosta, Patu, Bonder, Solernó, Willy y Pinasco; a Igna Ojea, Lucas Bali, Quimey y Laplagne, que influyeron de manera especial.

A mis compañeros en la docencia, en especial a Marcelo y Silvia.

A quienes me tuvieron como docente, por todo lo que aprendo de ellos.

A Juanma y Joel, que estuvieron todo el tiempo.

A mis compañeros de Computación, que formaron parte de una etapa clave y la hicieron memorable. A Lean, Ivan y Xavi.

A Mer y Miche; a Lula, Mili, Marianoide, Juancho y Mánol; que llenaron de alegría los primeros años.

A mis compañeros de militancia, al Conjunto Abierto. A Lupi y Santi, por meterme tantas ideas en la cabeza.

A todos mis compañeros del DM con los que compartí clases y juntadas de estudio, pero también viajes, juegos y conversaciones en los pasillos. Especialmente a mis amigos los vices.

A Luz, porque sin su ayuda el camino hasta acá hubiera sido mucho más difícil.

A Carlo, por su amistad y compañía en tantos momentos, incluso en esta tesis.

A mis padres, que me apoyaron siempre y son los autores de este logro personal. A mi familia que es un pilar fundamental.

A mi compañera de vida, porque su amor me hace quien soy todos los días.

Índice general

1. Elementos de la Teoría de Juegos: la forma normal **9**

1.1. Juegos no cooperativos y equilibrios 9

1.1.1. Definiciones generales y el concepto de equilibrio 9

1.1.2. Estrategias de seguridad 12

1.2. Juegos finitos de 2 jugadores: existencia de los Equilibrios 13

1.2.1. Juegos de suma cero y el teorema Minimax 14

1.2.2. Juegos de suma general y el teorema de Nash 16

1.3. Juegos jerárquicos finitos y equilibrios Stackelberg 18

1.3.1. Equilibrio Stackelberg en estrategias puras 19

1.4. Equilibrios Stackelberg en estrategias mixtas 23

2. La forma extensiva **27**

2.1. Información perfecta 27

2.1.1. Equilibrio de Nash 29

2.1.2. Equilibrios perfectos en subjuegos 30

2.2. Juegos de información imperfecta y el enfoque Bayesiano 33

2.2.1. Estrategias mixtas y conductuales 35

2.2.2. Estrategias y Creencias 37

2.2.3. Equilibrio secuencial 38

2.2.4. Juegos Bayesianos 41

3. Las ventajas de ser el líder **45**

3.1. ¿Extensivo o estratégico? 45

3.2. Estrategias inducibles 47
3.3. El equilibrio fuerte 56

4. Métodos de optimización para resolver juegos Stackelberg 59
4.1. Introducción y formulación inicial del problema 59
4.2. Distintas formulaciones de JSS para programación lineal entera mixta 62
4.2.1. Sobre la estructura de los pagos en aplicaciones de seguridad 64
4.3. Representación compacta 65
4.4. Modelos con incertidumbre observacional y de ejecución 68

5. La aplicación 75
5.1. Problemática 75
5.2. El modelo 77
5.3. El software 81

Introducción

La teoría de juegos podría llamarse teoría de las decisiones, como incluso algunos autores han propuesto. Modelar la interacción entre jugadores y razonar estratégicamente hasta encontrar la decisión óptima, en el amplio sentido de minimizar un costo o maximizar una utilidad: De esta manera podríamos resumir una idea general de lo que trata la teoría de juegos y la investigación operativa, y por lo tanto esta tesis. En el camino, nos encontramos con el desafío de ser fieles a la realidad que queremos modelar para lo cual las estructuras que utilicemos y las funciones que definan los pagos del juego son cruciales.

Por adecuarse a diferentes necesidades tanto teóricas como de aplicación la Teoría de Juegos ha demostrado ser un tema de estudio muy fructífero y prometedor, pues continúa estando presente en numerosas disciplinas: biología, economía, política, filosofía, computación, lógica. Los resultados obtenidos para las ciencias económicas han convertido a varios exponentes del área en premios Nobel: Nash, Harsanyi y Selten (1994), Schelling y Aumann (2005), Myerson, Hurwicks y Maskin (2007), Shapley y Roth (2012). Algunos, reconocidos aún fuera del ámbito científico.

Este trabajo, y la teoría que expondremos en él, tienen una meta: el desarrollo del motor matemático de un software que ayude a mejorar la seguridad del Aeropuerto de Ezeiza Ministro Pistarini, optimizando la disposición de los recursos humanos en los distintos sectores públicos. Para eso, revisaremos una pequeña porción del material bibliográfico con el que se cuenta, tanto en cuestiones básicas de la teoría de juegos no cooperativos como en los temas de investigación actuales en materia de investigación operativa aplicada en problemas de seguridad.

Comenzaremos con lo clásico: juegos *no cooperativos* en forma normal en el primer capítulo y forma extensiva en el segundo, los distintos conceptos de soluciones, equilibrios y estrategias puras, mixtas, conductuales. Expondremos la teoría desde el punto de vista de los juegos simultáneos y el equilibrio de Nash y luego estudiaremos la situación jerárquica líder-seguidor y el equilibrio Stackelberg. Haremos luego un rápido recorrido por los juegos con información imperfecta y nos

detendremos en el caso particular de los juegos Bayesianos, con los que alcanzaremos suficiente contexto teórico sobre el cual modelar y optimizar problemas con incertezas en los pagos.

En el tercer capítulo discutiremos brevemente sobre los puntos de vista utilizados en los dos primeros, y nos dedicaremos principalmente a las diferencias entre el juego simultáneo y el jerárquico, estableciendo el concepto de solución que se utiliza actualmente en los trabajos de optimización y de aplicación de los juegos Stackelberg para resolver problemas de seguridad, el equilibrio fuerte.

En el cuarto capítulo intentaremos esbozar un estado del arte sobre estos trabajos, centrándonos principalmente en los resultados obtenidos por el grupo de investigación de Fernando Ordóñez, codirector de esta tesis. En este sentido, plantearemos distintas formulaciones de los problemas discutiendo las diferencias entre ellas y sacaremos especial provecho de la estructura de los pagos cuando se trata de juegos de Seguridad Stackelberg. Confluiremos a modelos de programación lineal entera mixta, y alcanzaremos más profundidad con los casos en los que se tienen en cuenta incertezas observacionales y de ejecución.

El quinto y último capítulo está dedicado al trabajo realizado en el Aeropuerto de Ezeiza por un grupo bajo la dirección de Guillermo Durán y Fernando Ordóñez y compuesto por el Doctor en Computación Ivo Koch, Carlo Ferrari y el autor de esta tesis. Expondremos la problemática, el modelo que usamos para resolverla y especificaremos los tecnicismos sobre la arquitectura del software que está actualmente instalado y en etapa de prueba, a la espera de un análisis de impacto pertinente.

Capítulo 1

Elementos de la Teoría de Juegos: la forma normal

En este capítulo nos dedicaremos a la teoría de los juegos no cooperativos en los que el objetivo principal de los jugadores es optimizar sus pagos sin preocuparse por los correspondientes para sus adversarios. Introduciremos los conceptos básicos de la forma normal o estratégica, y definiremos las soluciones que son de nuestro interés. Nos basamos principalmente en el trabajo de Tamer Basar y Geert Jan Olsder [3].

1.1. Juegos no cooperativos y equilibrios

1.1.1. Definiciones generales y el concepto de equilibrio

Para dar una definición de juego no cooperativo debemos especificar distintos componentes que lo caracterizan:

- La cantidad de jugadores que participan,
- las acciones que puede tomar cada uno,
- la función objetivo que cada jugador quiere optimizar,
- el orden que hay en la elección de las estrategias y si este proceso se repite o no,
- la información de la que disponen los jugadores que pueda resultarles relevante a la hora de decidir y cómo ésta depende de las acciones tomadas

anteriormente en cada momento del tiempo en que se desarrolla el juego, y

- si existe un jugador (generalmente llamado "la naturaleza") cuyas acciones son resultado de algun proceso estocástico o *lotería* cuya distribución (fija) se conoce.

Consideraremos relevantes, por el momento, los primeros 3 ítems: los jugadores actúan una vez, el juego es estático y por lo tanto los jugadores no adquieren ninguna información de las acciones de los demás jugadores, y no participa la naturaleza.

Para ver las definiciones y conceptos más generales trabajaremos con un juego de N jugadores, con $\mathcal{N} := \{1, ..., N\}$ notando el conjunto de jugadores. La variable de acción o decisión del Jugador i será notada $x_i \in X_i$, donde X_i es el conjunto de estrategias del Jugador i. Este conjunto puede ser tanto finito como infinto, y en cuyo caso puede tener dimensión finita (por ejemplo el intervalo real unitario, $[0, 1]$) o infinita (por ejemplo las funciones continuas con soporte en el intervalo real unitario). Además, notaremos x al vector N-dimensional que contiene las estrategias de todos los jugadores $x := (x_1, ..., x_N)$. Si pensamos que podría haber restricciones para las acciones de los jugadores, entonces tiene sentido considerar $\Omega \subset X$ el conjunto de N-tuplas de variables de acciones factibles.

Si estamos en el caso en que los jugadores están minimizando, la función objetivo (de costo o pérdida) del Jugador i será $C_i(x_i, x_{-i})$, donde x_{-i} es el vector de acciones de todos los jugadores con excepción del i-ésimo. Si los jugadores están maximizando, entonces la función objetivo (de utilidad) del Jugador i será denotada $U_i(x_i, x_{-i})$. A lo largo de esta tesis usaremos tanto C_i como U_i, intercambiando entre ellas según la necesidad. Es importante observar que el juego en todos los jugadores minimizan sus costos C_i's, puede ser visto como uno en el que maximizan las funciones de utilidad $U_i = -C_i$.

Con esto queda definida la forma normal del juego, a la que típicamente asociaremos una matriz como función bilineal de pagos. La siguiente definición corresponde a nuestro concepto más basico de solución.

Definición 1.1. Un vector de estrategias $x^* \in \Omega$ es un *Equilibrio de Nash* (EN) si, $\forall i \in \mathcal{N}$, se tiene:

$$U_i(x_i^*, x_{-i}^*) \geq U_i(x_i, x_{-i}^*) \ \forall x_i \in X_i \text{ cumpliendo } (x_i, x_{-i}^*) \in \Omega,$$

o, en el caso en que los jugadores minimicen costos,

$$C_i(x_i^*, x_{-i}^*) \leq C_i(x_i, x_{-i}^*) \ \forall x_i \in X_i \text{ cumpliendo } (x_i, x_{-i}^*) \in \Omega.$$

Observemos que en el caso $N = 2$, si tenemos la particular situación en que $C_1 \equiv -C_2 =: C$, entonces el Juego se llama de Suma Cero (JSC), con el Jugador 1 minimizando C y el Jugador 2 maximizándola. En este caso, el equilibrio de Nash es equivalente a un *equilibrio de punto-silla* (EPS) que definiremos a continuación. Para eso, dejaremos de lado el conjunto de restricciones, es decir, asumimos que $X := X_1 \times X_2$, cualquier combinación de acciones es factible.

Definición 1.2. Un par de acciones $(x_1^*, x_2^*) \in X$ está en *equilibrio de punto-silla (EPS)* para un juego con función de utilidad C, si

$$C(x_1^*, x_2) \leq C(x_1^*, x_2^*) \leq C(x_1, x_2^*), \quad \forall (x_1, x_2) \in X$$

Esto además implica que podemos intercambiar el orden en que maximizamos y minimizamos inconsecuentemente, es decir:

$$\min_{x_1 \in X_1} \max_{x_2 \in X_2} C(x_1, x_2) = \max_{x_2 \in X_2} \min_{x_1 \in X_1} C(x_1, x_2) = C(x_1^*, x_2^*) =: C^*$$

donde la expresión de la izquierda se conoce como el *valor superior* del juego, la segunda expresión es el *valor inferior*, y C^* es conocido como el *valor* del juego. Estas definiciones se dan, en casos mas generales, con ínfimo y supremo en lugar de mínimo y maximo, respectivamente. Notemos que en general tenemos la desigualdad:

$$\min_{x_1 \in X_1} \max_{x_2 \in X_2} C(x_1, x_2) \geq \max_{x_2 \in X_2} \min_{x_1 \in X_1} C(x_1, x_2)$$

o más precisamente,

$$\inf_{x_1 \in X_1} \sup_{x_2 \in X_2} C(x_1, x_2) = \sup_{x_2 \in X_2} \inf_{x_1 \in X_1} C(x_1, x_2),$$

que se deduce directamente del hecho:

$$\sup_{x_2 \in X_2} C(x_1, x_2) \geq \inf_{x_1 \in X_1} C(x_1, x_2),$$

donde la expresión de la izquierda es una función que sólo depende de x_1 y la de la derecha sólo depende de x_2.

De esta manera, siempre que el valor del juego exista (o equivalentemente, siempre que exista un punto-silla), será único. En efecto, si $(\widehat{x_1}, \widehat{x_2})$ es otra solución punto-silla, se tiene que $C(\widehat{x_1}, \widehat{x_2}) = C^*$. Más aún, estas estrategias son intercambiables, es decir, $(\widehat{x_1}, x_2^*)$ y $(x_1^*, \widehat{x_2})$ también son EPS. Esta propiedad de los EPS no se extiende a múltiples equilibrios de Nash (para JSC): dos EN distintos, en general, implican valores distintos de las funciones de costo para los jugadores. Esto se debe a que el cambio unilateral de estrategia puede beneficiar

a un jugador pero implica un incremento en los costos de los otros, hecho que se desprende inmediatamente de la definición.

Así, si el JSC no tiene un valor, tenemos:

$$\min_{x_1 \in X_1} \max_{x_2 \in X_2} C(x_1, x_2) > \max_{x_2 \in X_2} \min_{x_1 \in X_1} C(x_1, x_2)$$

y por lo tanto no existe un punto-silla. En este caso diremos que *"el juego de suma cero no tiene un punto-silla en estrategias puras"*, lo cual nos incita a buscar un equilibrio en estrategias mixtas:

Definición 1.3. Una *estrategia mixta* es una distribución de probabilidad sobre el conjunto de acciones de cada jugador, a la cual denotaremos p_i para cada Jugador $i \in \mathcal{N}$.

En esta búsqueda, reemplazaremos C_i por el valor esperado en función de las estrategias mixtas elegidas por todos los jugadores y lo notaremos $J_i(p_1, \cdots, p_N)$. Llamaremos además $\mathcal{P}_i$ al conjunto de todas las distribuciones de probabilidad sobre X_i (y dejaremos de lado a Ω asumiendo que los conjuntos de restricciones son rectangulares). En el caso X_i finito, p_i será un vector aleatorio soportado en el simplex de probabilidad determinado por X_i.

Definición 1.4. Diremos que la N-tupla $(p_1^*, \cdots, p_N^*)$ está en *equilibrio de Nash para estrategias mixtas* (ENEM) si

$$J_i(p_i^*, p_{-i}^*) \leq J_i(p_i, p_{-i}^*) \quad \forall p_i \in \mathcal{P}_i \quad \forall i \in \mathcal{N}.$$

Esto nos conduce fácilmente, para el caso especial de los juegos de suma cero con $N = 2$, a la siguiente definición:

Definición 1.5. El par (p_1^*, p_2^*) constituye un *equilibrio de punto-silla para estrategias mixtas* (EPSEM) si

$$J(p_1^*, p_2) \leq J(p_1^*, p_2^*) \leq J(p_1, p_2^*) \quad \forall (p_1, p_2) \in \mathcal{P}$$

donde $J(p_1, p_2) = \mathbb{E}_{p_1, p_2}[C(x_1, x_2)]$ y $\mathcal{P} := \mathcal{P}_1 \times \mathcal{P}_2$ es el valor del juego de suma cero para estrategias mixtas.

1.1.2. Estrategias de seguridad

En los casos en los que no se tiene un equilibrio de Nash para estrategias puras, y los jugadores no necesariamente quieren adoptar estrategias mixtas, una alternativa es elegir la estrategia pura que mejor se adapte a un escenario pesimista,

es decir, que minimice los costos en el peor caso. Esto implicará que cada jugador escencialmente piense al juego como de suma cero, minimizando su función de costo contra la maximización colectiva de los demás jugadores. La estrategia (o acción, en este caso) que provee de un techo de pérdida para un jugador se llamará *estrategia de seguridad* para ese jugador, concepto que podemos definir formalmente, asumiendo que el producto de conjuntos de acciones es rectangular, como sigue:

Definición 1.6. Una estrategia $x_i^s \in X_i$ se dice *de seguridad* para el jugador i si cumple la relación:

$$\sup_{x_{-i}\in X_i} C_i(x_i^s, x_{-i}) = \inf_{x_i\in X_i} \sup_{x_{-i}\in X_{-i}} C_i(x_i, x_{-i}) =: \overline{C}_i$$

donde el supremo puede ser reemplazado por un máximo si el conjunto de acciones es finito.

Es importante observar que $\overline{C}_i$ es el valor superior del JSC que juega el Jugador i. Notemos también que incluso si las estrategias de seguridad para los jugadores, digamos $x^s := \{x_i^s,\ i \in \mathcal{N}\}$, son únicas, entonces la N-tupla no necesariamente es un equilibrio de algún tipo.

Esta noción de estrategia de seguridad puede extenderse naturalmente a estrategias mixtas usando la notación introducida previamente:

Definición 1.7. Diremos que $p_i^s \in \mathcal{P}_i$ sera una *estrategia de seguridad mixta* para el Jugador i si

$$\sup_{p_{-i}\in \mathcal{P}_i} J_i(p_i^s, p_{-i}) = \inf_{p_i\in \mathcal{P}_i} \sup_{p_{-i}\in \mathcal{P}_{-i}} J_i(p_i, p_{-i}) =: \overline{J}_i$$

Observación 1. Si el juego original era un juego de suma cero con 2 jugadores, y los valores superior e inferior son iguales, entonces las estrategias de seguridad para los jugadores deben estar en EPS. Si estos valores no son iguales en estrategias puras, pero lo son en mixtas, entonces las estrategias de seguridad mixtas para los jugadores deben estar en EPSEM.

1.2. Juegos finitos de 2 jugadores: existencia de los Equilibrios

En esta sección nos centraremos en el caso particular en el que intervienen 2 agentes en el proceso de decisión: dos jugadores que notaremos **J**1 y **J**2. Trabajaremos las situaciones en las que los conjuntos de estrategias de los jugadores son

finitos y analizaremos la existencia de los equilibrios y cómo encontrarlos cuando existan. Ejemplificaremos algunas situaciones y daremos un marco notacional sobre el cual trabajar posteriormente. Aprovecharemos el hecho de que tenemos sólo dos jugadores para simplificar levemente la notación que usamos hasta aquí.

1.2.1. Juegos de suma cero y el teorema Minimax

Un juego se dirá de suma cero cuando los costos que se le generan a un jugador son iguales a la ganancia del otro. Para un juego de este tipo debemos especificar la cantidad de elementos que poseen los conjuntos de acciones, que notaremos $x \in X$ e $y \in Y$ para **J**1 y **J**2 respectivamente, y la función objetivo $C(x, y)$ definida sobre el producto de estos conjuntos finitos. Convengamos que **J**1 minimiza sus costos y **J**2 maximiza sus ganancias, recordando que como el juego es de suma cero tenemos que $C(x, y) = C_1(x, y) = U_2(x, y) = -C_2(x, y)$. Llamemos $n = |X|$ y $m = |Y|$ la cantidad de acciones posibles que pueden tomar **J**1 y **J**2 respectivamente, y supongamos que éstas acciones están ordenadas bajo cierta convención (que puede ser arbitraria).

Entonces, podemos asociar nuestro juego con una matriz A de dimensión $n \times m$ cuyas entradas corresponden con la función de costos $C(x, y)$, siguiendo el mismo orden que tienen los conjuntos de acciones, es decir, $A_{ij} = C(x, y)$ cuando x es el i-ésimo elemento de X y y el j-ésimo elemento de Y. Las elecciones del Jugador 1 son entonces las filas de la matriz y las del Jugador 2 sus columnas.

Ejemplo. Es un ejemplo clásico en el que ambos jugadores eligen entre dos opciones. Si coinciden entonces gana el Jugador 1 y si no gana el Jugador 2. Podemos asociarle la matriz:

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

En este ejemplo no hay ninguna combinación fila-columna para la cual los jugadores no tengan un incentivo de desviarse y mejorar sus pagos.

La pregunta que arroja este ejemplo es cuándo existe un equilibrio de punto silla para estrategias mixtas. Supongamos que ahora **J**1 decide elegir fila 1 y fila 2 con probabilidad 1/2, sin importar si **J**2 juega columna 1 o 2, su costo esperado será de 0. Luego, en respuesta a la elección equiprobable de **J**1, el Jugador 2 es indeferente entre sus posibles acciones: puede elegir jugar columna 1, columna 2, o cualquier probabilidad mixta entre ambas. Razonamos de la misma manera si pensamos en que **J**2 juega una estrategia mixta $(1/2, 1/2)$, esta vez para **J**1 el costo que tendrá por resultado será 0 independientemente de qué estrategia elija.

Así, el par $(p^* = (1/2, 1/2), q^* = (1/2, 1/2))$ será un EPSEM con valor asociado 0, y de hecho es el único.

Para formalizar esta situación, sea A una matriz de dimensión $n \times m$ representando el juego de suma cero, y como antes, p y q dos vectores de probabilidad para $\mathbf{J}1$ y $\mathbf{J}2$ respectivamente (que consideraremos como columnas de longitud n y m resp. cuyas coordenadas estan en el intervalo real unitario y suman uno). Podemos escribir entonces la función de costo esperado como

$$J(p, q) := p^t Aq.$$

La existencia de los equilibrios de punto silla en estrategias mixtas queda resuelta en 1982 con el Teorema Minimax de John Von Neumann [28].

Teorema 1.8. *Todo juego finito de suma cero para dos jugadores admite un punto silla en estrategias mixtas, es decir, dada A la matriz asociada al juego, existe un par (p^*, q^*) tal que para cualquier otro par de vectores de probabilidad p y q se tiene que*

$$p^{*t} Aq \leq p^{*t} Aq^* \leq p^t Aq^*$$

donde $p^{*t} Aq^*$ es el valor del juego *en estrategias mixtas.*

Existen distintas demostraciones de este teorema, que están basadas en razonar por el absurdo vía inducción en el tamaño de la matriz de pago, en el teorema de punto fijo de Brouwer o en el Corolario de Hahn-Banach del hiperplano separador. Escribiremos aquí una demostración basada en el teorema de dualidad para programación lineal que podemos encontrar en [8].

Demostración. Consideremos el problema desde la mirada del Jugador 1. Estamos buscando $p_1, ..., p_n$ que maximice

$$\min_{1 \leq j \leq m} \sum_{i=1}^{n} p_i A_{ij} = \min_{q \in \mathcal{Q}} p^t Aq$$

sujeto a que $p \in \mathcal{P}$, esto nos deja el problema:

$$\begin{aligned} \max_{p} \quad & \min_{1 \leq j \leq m} \sum_{i=1}^{n} p_i A_{ij} \\ s.a. \quad & \sum_{i=1}^{n} p_i = 1, p \geq 0 \end{aligned}$$

El inconveniente es que la función objetivo no es lineal y por lo tanto apelaremos al procedimiento habitual de linearizarla agregando una variable adicional v restringiéndola a que sea menor o igual que el mínimo en la función objetivo e intentando maximizarla. El problema ahora queda:

$$\begin{aligned} \max_{p,v} \quad & v \\ s.a. \quad & v \leq \sum_{i=1}^{n} p_i A_{ij} \quad \forall 1 \leq j \leq m \\ & \sum_{i=1}^{n} p_i = 1, p \geq 0 \end{aligned}$$

resultando un problema de programación lineal, cuya solución explícita está a nuestra disposición vía el algoritmo *SIMPLEX*. Para profundizar en el tema se puede consultar [6].

Si razonamos de manera análoga, pero esta vez desde el lugar del Jugador 2 entonces obtenemos el problema:

$$\begin{aligned} \min_{q,w} \quad & w \\ s.a. \quad & w \geq \sum_{j=1}^{m} A_{ij} q_j \quad \forall 1 \leq i \leq n \\ & \sum_{j=1}^{m} q_j = 1, q \geq 0 \end{aligned}$$

En este punto es crucial notar que el problema planteado para **J**2 es el dual del planteado para **J**1, y por lo tanto, por el teorema de dualidad [6], el valor óptimo de ambos coincide y es $p^* A q^*$, como queríamos. □

La demostración además nos indica un mecanismo mediante el cual obtener la solución. Estas estrategias son también denominadas *estrategias minimax* por razones evidentes.

1.2.2. Juegos de suma general y el teorema de Nash

Considereramos ahora los juegos de suma general (no necesariamente cero) para N Jugadores como al comienzo y estudiaremos la existencia de los equilibrios de

Nash que introdujimos anteriormente. Es fácil encontrar ejemplos para los cuales no existen dichos equilibrios en estrategias puras, la pregunta es entonces cuándo existen EN en estrategias mixtas. Esto fue resuelto en 1951 por John Nash (ver [18]) y está plasmado en el siguiente teorema:

Teorema 1.9 (Teorema de Nash)**.** *Todo juego de suma general para N jugadores con conjuntos de acciones finitos admite un equilibrio Nash en estrategias mixtas.*

Nuevamente, podemos encontrar numerosas demostraciones, con diferentes hipótesis y tecnicismos. Esta vez nos remitiremos a la propuesta por John Geanakoplos, que podemos encontrar en [11]. La demostración que expondremos aquí presupone que la función de utilidad es cóncava y se basa en el Teorema del Punto Fijo de Brauwer y en el siguiente Lema de la Permutación Cóncava:

Lema 1.10. *Sea $X \subset \mathbb{R}^n$ un conjunto convexo y sea $x \in X$. Sea $J : X \to \mathbb{R}$ una función cóncava. Entonces* $\arg\max_{y\in X} [J(y) - \|x - y\|^2]$ *tiene al menos un elemento y si* $x = \arg\max_{y\in X} [J(y) - \|x - y\|^2]$*, entonces* $x \in \arg\max_{y\in X} J(y)$*.*

Demostración. Como J es cóncava en y y $-\|y - x\|^2$ es estrictamente cóncava en y entonces $[J(y) - \|x - y\|^2]$ es estrictamente cóncava y por lo tanto,
$\arg\max_{y\in X} [J(y) - \|x - y\|^2]$ no puede contener dos puntos distintos. Supongamos ahora que $x = \arg\max_{y\in X} [J(y) - \|x - y\|^2]$ y tomemos $y \in X$. Como X es convexo y J cóncava, para cualquier $\varepsilon \in (0, 1)$,

$$\begin{aligned} 0 &\geq \{J((1-\varepsilon)x + \varepsilon y) - \|(1-\varepsilon)x + \varepsilon y - x\|^2\} - \{J(x) - \|x - x\|^2\} \\ &= J((1-\varepsilon)x + \varepsilon y) - \varepsilon^2\|y - x\|^2 - J(x) \\ &\geq (1-\varepsilon)J(x) + \varepsilon J(y) - J(x) - \varepsilon^2\|y - x\|^2 \\ &= \varepsilon(J(y) - J(x)) - \varepsilon^2\|y - x\|^2 \end{aligned}$$

y por lo tanto

$$J(y) - J(x) \leq \varepsilon\|y - x\|^2 \text{ para todo } \varepsilon > 0$$

Luego,

$$J(x) \geq J(y) \text{ para todo } y \in X$$

. □

Demostración del teorema de Nash. Consideremos un juego de N jugadores $\Gamma = (\mathcal{N}; \mathcal{P}_1, ..., \mathcal{P}_N; J_1, ..., J_N)$ y supongamos que para todo $p_{-n} :=$
$(p_1, ..., p_{n-1}, p_{n+1}, ..., p_N) \in X_{-n}$ con $X_{-n} := X_1 \times ... \times X_{n-1} \times X_{n+1} \times ... \times X_N$ se tiene que $J_n(p_n, p_{-n})$ es una función cóncava en p_n. Con esta notación, $p^* \in \mathcal{P}$

será un equilibrio de Nash si cumple que para todo $n \in \mathcal{N}$ y todo $p_n \in \mathcal{P}_n$, $J_n(p^*) \geq J_n(p_n, p^*_{-n})$.

Definamos la función $\varphi : \mathcal{P} \to \mathcal{P}_n$ como sigue:

$$\varphi_n(p_1, ..., p_n, ..., p_N) = \arg \max_{\overline{p}_n \in \mathcal{P}_n} \left[J_n(\overline{p}_n, p_{-n}) - \|\overline{p}_n - p_n\|^2 \right].$$

Observemos que la función a la que le buscamos el argumento máximo es suma de una funcion continua, cóncava en p_n y otra cuadrática y negativa, y por lo tanto se trata de una función continua estrictamente cóncava. Como $\mathcal{P}_n$ es compacto y convexo, φ está bien definida. Más aún, el maximando es continuo como función de $p = (p_1, ..., p_N)$, por principio del máximo, φ_n resulta una función continua.

Consideremos ahora a la función $\varphi : X \to X$ definida como $\varphi = (\varphi_1, ..., \varphi_N)$. Claramente, φ resulta continua, y por el Teorema de punto fijo de Brouwer, existe un p^* tal que $\varphi(p^*) = p^*$.

Ahora bien, por el Lema de la Perturbación Cóncava, tenemos que para cualquier estrategia $p_n \in \mathcal{P}_n$, $J_n(p_n, p^*_{-n}) \leq J_n(p^*)$ y por lo tanto p^* es un equilibrio de Nash. □

1.3. Juegos jerárquicos finitos y equilibrios Stackelberg

El equilibrio de Nash como concepto de solución que analizamos anteriormente provee de un resultado razonable para juegos no cooperativos cuando los roles de los jugadores son simétricos, es decir, cuando ningún jugador domina el proceso de decisión. Sin embargo, hay otro tipo de situaciones en la que intervienen agentes que toman decisiones de manera no cooperativa en las que alguno de los jugadores tiene la capacidad de forzar la estrategia de los demás jugadores mediante el anuncio de su estrategia, y para este tipo de problemas de decisión podemos desarrollar una teoría similar pero sustancialmente diferente.

En 1934, Heinrich Freiherr von Stackelberg (1905-1943) describió en su obra [29] un modelo de liderazgo para situaciones de oligopolio económico (o duopolio en el caso $N = 2$) diferenciando dos roles disparejos en la interacción de los agentes: el jugador que está en esta posición más importante en el proceso de toma de decisión será llamado el líder, y los otros jugadores que responden racionalmente a la estrategia del líder serán los seguidores.

En éste camino, se pueden considerar muchos niveles de jerarquía, con varios líderes y seguidores, pero para introducir los conceptos de manera clara analizare-

mos aquí los problemas de decisión jerárquica que involucran solo dos jugadores: un líder y un seguidor.

1.3.1. Equilibrio Stackelberg en estrategias puras

Para introducir el concepto de solución jerarquica de equilibrio de Stackelberg, consideremos el juego bimatricial (A, B) dado por :

$$A = \begin{bmatrix} 0 & 2 & 3/2 \\ 1 & 1 & 3 \\ -1 & 2 & 2 \end{bmatrix}, B = \begin{bmatrix} -1 & 1 & -2/3 \\ 2 & 0 & 1 \\ 0 & 1 & -1/2 \end{bmatrix}$$

Donde el conjunto de acciones para ambos jugadores está compuesto por Izquierda (I), Medio (M) y Derecha (D), y veremos la estrategia del Jugador 1 en las filas y la del Jugador 2 en las columnas. (nuestra convencion estandar).

Después de observar con detenimiento los costos del juego, concluimos que admite un único EN en estrategias puras, que es $\{M, M\}$, con costos resultantes $(1, 0)$. Ahora bien, ¿qué sucedería si el Jugador 1 anunciara su estrategia al Jugador 2? Esto es pensar que los roles de ambos jugadores no son simétricos y el Jugador 1 actúa como el líder en el proceso de decisión. Así, si el Jugador 2, el seguidor, conoce lo que jugará su contrincante simplemente debe elegir la acción que le signifique el costo mínimo.

Una vez entendido esto, el Jugador 1, antes de anunciar su estrategia, deberá elegir la acción que minimice sus costos considerando la respuesta racional que su elección implicará. De esta manera, si **J**1 elige I, entonces **J**2 tiene una única respuesta que minimiza sus costos que es I, que significa un costo de 0 para **J**1. Si el líder elige M, entonces la respuesta óptima de **J**2 también es única (que es M) y el costo que implica en **J**1 es 1. Por último, si **J**1 elige D, entonces **J**2 eligirá D y el costo para el líder asociado a esta elección será 2. Analizada la situación, el menor costo para **J**1 es 0 y la estrategia óptima I.

Diremos entonces que I es la estrategia Stackelberg para el líder (**J**1) y el par $\{I, I\}$ es la solución Stackelberg para el juego (A, B) con **J**1 en el rol de líder. Además, el par de costos $(0, 1)$ es el valor (equilibrio) Stackelberg para el juego con **J**1 en el rol de líder. Es interesante notar que este valor es estrictamente mejor que el valor del equilibrio de Nash para ambos jugadores, aunque no vale en general. Si ahora consideramos a **J**2 en el rol de líder, entonces el equilibrio Stackelberg es $\{I, D\}$ con valores $(3/2, -2/3)$ que es claramente desfavorable para **J**1 (el seguidor) cuando se compara con el costo del EN. De todas maneras, para **J**2 (el líder), el valor Stackelberg es nuevamente mejor que el del EN (menor costo).

El concepto de equilibrio Stackelberg (ES) que introdujimos recién mediante juegos bimatriciales es aplicable a cualquier juego finito de dos jugadores en forma normal aunque hicimos uso de un hecho que se da en ese ejemplo y que no sucede en general: la respuesta óptima del seguidor a cualquier estrategia del líder era única. Si este requerimiento no se satisface, entonces puede haber ambigüedad en cuál es la mejor respuesta para el seguidor y por lo tanto en los posibles costos que debe considerar el líder a la hora de elegir su estrategia óptima. Para ilustrar esto consideremos el siguiente ejemplo de juego bimatricial:

$$A = \begin{bmatrix} 0 & 1 & 3 \\ 2 & 2 & -1 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & -1 \end{bmatrix}$$

donde las posibles acciones del líder **J**1 son las filas $\{I, D\}$ y las acciones del seguidor son las columnas $\{I, M, D\}$. En este caso, si **J**1 elige (y anuncia) I entonces **J**2 puede jugar óptimamente tanto I como M y sus costos asociados serán 0 en ambos casos. Si **J**1 elige D, **J**2 nuevamente tiene dos acciones que minimizan sus costos, I y D. Esta multiplicidad en respuestas óptimas para el seguidor se traduce en diferentes costos para el líder, y por lo tanto no podemos aplicar directamente el concepto de solución que presentamos anteriormente. Sin embargo, podemos asumir que el líder optará por la estrategia que minimice sus costos en el peor de los casos, reduciendo los riesgos. Luego, para este modo de juego, los costos asegurados para **J**1 son 1 y 2 para I y R respectivamente. Diremos entonces que $x^* = I$ es la única estrategia óptima Stackelberg para el líder. El costo correspondiente para **J**1 será C_1^*. Debemos notar que la multiplicidad de respuestas óptimas para **J**2 puede implicar que el costo de **J**1 sea en realidad menor, dependiendo de la elección del seguidor, cuyas estrategias óptimas son $y = I$ y $y = M$. De esta manera, los valores del juego pueden ser tanto $(1, 0)$ como $(0, 0)$, y por lo tanto no podemos hablar de unicidad del valor final para el equilibrio Stackelberg. Aún así, es interesante notar que el EN del juego bimatricial propuesto es $\{D, D\}$ con valor $(-1, -1)$ que es más favorable para ambos que el equilibrio Stackelberg.

Formalizaremos ahora los conceptos introducidos anteriormente con juegos bimatriciales para juegos de dos jugadores finitos teniendo en consideración los conjuntos X y Y los espacios de estrategias puras para **J**1 y **J**2 respectivamente, y $C_i(x, y)$ notando el costo para el Jugador i correspondiente al par $(x \in X, y \in Y)$.

Definición 1.11. En un juego finito para dos jugadores, el conjunto $R_2(x) \subset Y$ definido para cada $x \in X$ como

$$R_2(x) = \{z \in Y : C_2(x, z) \leq C_2(x, y) \forall y \in Y\}$$

es el *conjunto de respuestas óptimas* (o reacción racional) de **J**2 para la estrategia $x \in X$ de **J**1.

Definición 1.12. En un juego finito para dos jugadores con **J**1 en el rol del líder, una estrategia $x^* \in X$ se llama una *estrategia de equilibrio Stackelberg* para el líder, si cumple:

$$\max_{y \in R_2(x^*)} C_1(x^*, y) = \min_{x \in X} \max_{y \in R_2(x)} C_1(x, y) =: C_1^*.$$

El valor C_1^* es el costo de Stackelberg para el líder. Si en cambio, **J**2 es el líder la definición es equivalente intercambiando los subíndices 1 y 2.

Teorema 1.13. *Todo juego finito para dos jugadores admite una estrategia de equilibrio Stackelberg para el líder.*

Demostración. Tanto X como Y son conjuntos finitos, y por lo tanto lo es $R_2(x)$ para cada x. La existencia de la estrategia resulta evidente pues los conjuntos donde maximizamos y minimizamos son finitos. □

Observación 2. La estrategia Stackelberg para el seguidor no necesariamente es única como vimos anteriormente, pero la no unicidad de estrategias de equilibrio no presupone un problema (como sucedía con los equilibrios de Nash) dado que el costo de Stackelberg para el líder es único.

Observación 3. Si para cada $x \in X$, $R_2(x)$ consta de un único elemento, entonces existe una biyección $T_2 : X \to Y$ tal que $y \in R_2(x)$ implica $y = T_2(x)$. Esto se corresponde con las situaciones en las que la respuesta óptima del seguidor es única sea cual sea la estrategia que juegue el líder y esto concluye en una versión simplificada de la definición de equilibrio:

$$C_1(x^*, T_2(x^*)) = \min_{x \in X} C_1(x, T_2(x)) =: C_1^*$$

Desde el punto de vista del seguidor, la estrategia de equilibrio es cualquier respuesta óptima a la estrategia de Stackelberg anunciada por el líder. Más precisamente,

Definición 1.14. Sea $x^* \in X$ una estrategia Stackelberg para el líder (**J**1). Entonces, cualquier elemento $y^* \in R_2(x^*)$ es una *estrategia óptima* para el seguidor (**J**2) que está *en equilibrio* con x^*. El par $\{x^*, y^*\}$ es una *solución Stackelberg* para el juego con **J**1 en rol de líder, y par de costos $(C_1(x^*, y^*), C_2(x^*, y^*))$ es su correspondiente *valor de equilibrio Stackelberg*.

Observación 4. En la definición anterior, el valor $C_1(x^*, y^*)$ puede ser menor que el costo de Stackelberg C_1^* (algo que notamos anteriormente en el juego bimatricial). De todas maneras, si $R_2(x^*)$ consta de un único elemento, entonces estos valores de costo coinciden.

Consideremos un juego finito para dos jugadores, notemos nuevamente C_1^* al costo de Stackelberg para el líder (**J**1) y C_1^N algún costo proveniente de un equilibrio de Nash para el mismo jugador. Como pudimos observar en ejemplos anteriores, no necesariamente sucede que C_1^* es menor que C_1^N, en particular, cuando la respuesta óptima del seguidor no es única. La siguiente proposición nos provee de una condición suficiente bajo la cual el líder no actúa "peor" en el juego jerárquico que en el juego simultáneo:

Proposición 1.15. *Sea un juego finito para dos jugadores, y sean C_1^* y C_1^N como definimos recién. Si $\forall x \in X$ se tiene que $R_2(x)$ consta de un único elemento, entonces*

$$C_1^* \leq C_1^N$$

Demostración. Supongamos que existe un equilibrio de Nash $\{x', y'\}$ cuyo costo correspondiente para **J**1 es menor que el valor Stackelberg, es decir,

$$C_1^* > C_1(x', y').$$

Como $R_2(x)$ tiene un único elemento, sea $T_2 : Y \to X$ la biyección introducida en 3. Luego, $x' = T_2(y')$ y reemplazando esto en la expresión anterior obtenemos

$$\min_{x \in X} C_1(x, T_2(x)) = C_1^* > C_1(x', y')$$

concluyendo un absurdo. □

Observación 5. Uno está tentado a pensar que si un juego de suma general admite un único EN y una única strategia Stackelberg (x^*) para el líder, y además $R_2(x^*)$ tiene un solo elemento, entonces la desigualdad de la Proposición 1.15 debería seguir valiendo. Sin embargo, esto no es cierto como podemos ver con el siguiente ejemplo de juego bimatricial.

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix}, \; B = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$$

Como podemos verificar facilmente, existe un único equilibrio de Nash $\{D, D\}$ y una única estrategia de Stackelberg $x^* = I$ para el líder **J**1. Más aún, la respuesta óptima del seguidor también es única: $y^* = I$. Sin embargo, $0 = C_1^* > C_1^N = -1$. Este contraejemplo muestra que nuestra hipótesis suficiente no puede ser relajada de manera satisfactoria.

1.4. Equilibrios Stackelberg en estrategias mixtas

Cuando trabajamos con equilibrios de Nash nos encontramos con que no siempre existe el equilibrio en estrategias puras, y vimos la necesidad de introducir las estrategias mixtas y concluimos la existencia de equilibrios no cooperativos. Como vimos en el Teorema 1.13 siempre es posible encontrar soluciones de equilibrio puras para los juegos jerárquicos de tipo Stackelberg y en principio no tendríamos necesidad de introducir estrategias mixtas. Además, como el líder anuncia su estrategia al seguidor, parece poco razonable imaginar que este anuncio sea una estrategia mixta. Sin embargo, como ilustraremos en el siguiente ejemplo, hay situaciones en las que el uso de estrategias mixtas puede reducir los costos del líder si se razona adecuadamente.

Ejemplo. Consideremos el juego bimatricial (A, B) para $\mathbf{J}1$ actuando como líder con estrategias puras $\{I, D\}$ representadas en las filas de las matrices:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1/2 & 1 \\ 1 & 1/2 \end{bmatrix}$$

Este juego admite dos soluciones de equilibrio Stackelberg en estrategias puras: II y DD. El valor Stackelberg en ambos casos es $(1, 1/2)$. Sin embargo, si $\mathbf{J}1$ adopta la estrategia mixta de jugar I y D con probabilidad $1/2$ cada una, el costo promedio que obtendrá será de $1/2$, independientemente de la estrategia que juegue $\mathbf{J}2$. Este valor $J_1 = 1/2$ es claramente menor que el costo Stackelberg del líder en estrategias puras, y puede demostrarse que es el único costo Stackelberg para el líder en estrategias mixtas, puesto que cualquier desviación de $(1/2, 1/2)$ significa un aumento en los costos para el líder si tenemos en cuenta la respuesta óptima del seguidor.

El resultado anterior refleja la importancia que tiene estudiar equilibrios en estrategias mixtas en juegos de suma general para dos jugadores y muestra que es posible encontrar una estrategia que reduzca los costos para el líder. Consideraremos entonces este tipo de juegos en su forma normal y lo asociaremos con el juego bimatricial (A, B). En concordancia con lo descrito en secciones anteriores, notaremos $\mathcal{P}_i$ al espacio de estrategias mixtas de $\mathbf{J}i$, a los elementos $p_i \in \mathcal{P}_i$ y J_i a los costos esperados, $i = 1, 2$.

Definición 1.16. Para un juego bimatricial (A, B), el conjunto $RM_2(p) = \{q' \in \mathcal{Q} : p^t B q' \leq p^t B q \ , \ \forall q \in \mathcal{Q}\}$ es el conjunto de respuesta óptima (reacción racional) de $\mathbf{J}2$ en estrategias mixtas, para la estrategia mixta $p \in \mathcal{P}$ de $\mathbf{J}1$.

Definición 1.17. En un juego bimatrical (A, B) con **J**1 en el rol de líder, una estrategia mixta $p^* \in \mathcal{P}$ se llama una *estrategia mixta de equilibrio Stackelberg* para el líder si se cumple:

$$\max_{q \in RM_2(p^*)} p^{*t} A q = \inf_{p \in \mathcal{P}} \max_{q \in RM_2(p)} p^t A q =: J_1^*.$$

El valor J_1^* es el *costo de Stackelberg* en estrategias mixtas para el líder.

Aquí es importante notar que el "máximo" en la definición anterior siempre existe, pues para cada $p \in \mathcal{P}$, $p^t A q$ es continua en q, y $RM_2(p)$ es un subconjunto cerrado y acotado de $\mathcal{Q}$ (que es un simplex de dimensión finita). Por otro lado, el ínfimo no siempre podrá cambiarse por un mínimo, a menos que el problema admita una estrategia de equilibrio Stackelberg para el líder. El siguiente es un ejemplo de esta no existencia.

Ejemplo. Consideremos una version modificada del juego bimatricial del ejemplo anterior.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; B = \begin{bmatrix} 1/2 & 1 \\ 1 & 1/3 \end{bmatrix}$$

Con **J**1 en el rol del líder, este juego también admite dos estrategias puras de equilibrio Stackelberg, que son $\{I, I\}$ y $\{D, D\}$ y el costo de Stackelberg para el líder es $C_1^* = 1$. Ahora bien, si el líder utiliza la estrategia $p = [p, 1-p]$ podemos calcular que

$$J_2(p, q) = p^t B q = (\frac{-7}{6} p + \frac{2}{3}) q + \frac{2}{3} p + \frac{1}{3},$$

donde $q = [q, 1-q]$ denota la estrategia mixta de **J**2.

Entonces, podemos ver que la estrategia del líder $p = 4/7$ hace constante el pago del seguidor, dejándolo indiferente y determinando el conjunto de respuesta óptima de **J**2 de la siguiente manera:

$$R_2(p) = \begin{cases} \{q = (1, 0)\} & \text{si } p > 4/7 \\ \{q = (0, 1)\} & \text{si } p < 4/7 \\ \mathcal{Q} & \text{si } p = 4/7 \end{cases}$$

De esta manera, para $p > 4/7$ el seguidor elige D con probabilidad 1, y esto significa un costo promedio de $J_1 = p$ para **J**1. Para $p < 4/7$ en cambio, **J**2 elige I con probabilidad 1 y esto se traduce en un costo $J_1 = 1 - p$ para **J**1. Claramente, el líder preferirá quedarse en esta última región cuyo costo es menor. De hecho, si **J**1 juega $p = (4/7 - \varepsilon, 3/7 + \varepsilon)$ para ε suficientemente pequeño, su costo será de $J_1 = 3/7 + \varepsilon$, pues **J**2 jugará su única respuesta óptima que es la estrategia

pura $y = D$. Como $\varepsilon > 0$ puede tomarse arbitrariamente pequeño, llegamos a la conclusión de que el costo Stackelberg en estrategias mixtas será, tomando ínfimo, $J_1^* = \frac{3}{7}$ que es estrictamente menor que $C_1^* = 1$. Dicho esto, el líder no tiene una estrategia mixta que esté en equilibrio Stackelberg, puesto que para nuestro único candidato $p' = (4/7, 3/7)$, el conjunto de respuesta óptima es todo $\mathcal{Q}$ y por lo tanto $\max_{q \in \mathcal{Q}} p'^t Aq = 4/7$ que es mayor que J_1^*.

El ejemplo anterior es una muestra de que no siempre existe una solución de equilibrio Stackelberg cuando consideramos estrategias mixtas, pero a pesar de esto, podemos encontrar en ellas una reducción de los costos, empleando estrategias no necesariamente óptimas, por ejemplo $p = (4/7 - \varepsilon, 3/7 + \varepsilon)$. De hecho, siempre que se mantenga una desigualdad estricta $J_1^* < C_1^*$ podremos encontrar una estrategia mixta aproximada para el líder que mantenga los costos menores que si sólo consideramos estrategias puras. Si ambos valores son iguales, será razonable utilizar la estrategia de Stackelberg pura, que siempre existe como hemos visto en el Teorema 1.13. La siguiente proposición nos asegura que sólo puede pasar alguna de las dos situaciones recién mencionadas.

Proposición 1.18. *Para todo juego finito de dos jugadores se tiene que*

$$J_1^* \leq C_1^* \tag{1.1}$$

Demostración. Notemos P al subconjunto de $\mathcal{P}$ que consiste en todas las distribuciones de un solo punto y análogamente $Q \subset \mathcal{Q}$. Notemos que P es equivalente a X y Q a Y. Luego, para cada $p \in P$,

$$\min_{q^t \in \mathcal{Q}} pBq = \min_{q^t \in Q} pBq$$

pues cualquier solución que minimice puede ser reemplazada por un elemento de Q. Esto implica que, para cada $p \in P$, los elementos de $RM_2(p)$ son distribuciones de probabilidad en $R_2(p)$, donde este último conjunto podría ser definido como en 1.16 reemplazando $\mathcal{Q}$ por Q. Ahora bien, como $P \subset \mathcal{P}$,

$$J_1^* = \min_{p \in \mathcal{P}} \max_{q \in RM_2(p)} p^t Aq \leq \min_{p \in P} \max_{q \in RM_2(p)} p^t Aq,$$

y más aún, debido a la relacion mencionada entre RM_2 y R_2, el último valor es igual a

$$\min_{p \in P} \max_{q \in R_2(p)} p^t Aq$$

que por definición es es C_1^*, dado que $R_2(p)$ es equivalente al conjunto de respuesta óptima en estrategias puras del seguidor, como definimos en 1.11. Luego, $J_1^* \leq C_1^*$, como queríamos. □

El cómputo de las soluciones de equilibrio Stackelberg en estrategias mixtas no es tan inmediato como lo es cuando trabajamos en estrategias puras, pues los espacios $\mathcal{P}_i$ ya no son finitos. La técnica estándar es determinar soluciones que minimicen $p^t Bq$ como funciones de p, lo que nos conduce a descomponer $\mathcal{P}$ en regiones sobre las cuales podemos determinar los conjuntos de respuesta óptima, como ejemplificamos anteriormente. Una vez hecho esto nos resta minimizar $p^t Aq$ sobre p, sujeto a las restricciones impuestas por los conjuntos de reacción racional. Éste procedimiento nos provee además de estrategias aproximadas al óptimo para el líder, cuando no existe el equilibrio Stackelberg en estrategias mixtas (ni el costo Stackelberg).

Capítulo 2

La forma extensiva

Hay otras maneras de representar un juego además de la forma normal o estratégica, que como vimos es compacta y útil para un análisis inmediato (al menos en principio). Una de ellas es la forma extensiva, que plantea un seguimiento del proceso de toma de decisiones, describiendo de manera secuencial los problemas de decisión que encuentra cada jugador en la situación estratégica que se tiene en consideración. Esta forma de representación tiene en cuenta el desarrollo de un juego: las nociones de posición y movimiento.

Nos encontraremos con que los equilibrios de Nash no son las mejores soluciones para estos modelos, dado que no tienen en cuenta esta estructura secuencial, y definiremos entonces la noción alternativa de *equilibrios perfectos en subjuegos*, en los cuales los jugadores reevalúan sus planes a medida que el juego se desarrolla. Nos encontraremos con que los equilibrios Stackelberg son casos particulares de éstos y analizaremos su existencia para estrategias mixtas en este contexto.

2.1. Información perfecta

Diremos que los jugadores tienen información perfecta si en cada momento en que un jugador debe tomar una decisión tiene pleno conocimiento de las decisiones que se tomaron y los eventos que se sucedieron anteriormente en el juego. Inicialmente, y por simplicidad, nos enfocaremos en los juegos en los que no hay dos jugadores que tomen decisiones al mismo tiempo y en los que no hay intervención de la naturaleza (no hay aleatoriedad).

Definición 2.1. Un *Juego extensivo con información perfecta* consta de las siguientes componentes:

- Un conjunto $\mathcal{N}$ (el conjunto de jugadores).
- Un conjunto de secuencias H (finito o no) que satisface:
  - La secuencia vacía $\varnothing$ pertenece a H.
  - Si $(a_k)_{k=1,\dots,K} \in H$ (donde K podría ser infinito) y $L \leq K$ entonces $(a_k)_{k=1,\dots,L} \in H$.
  - Si $(a_k)_{k=1}^{\infty}$ cumple que $(a_k)_{k=1,\dots,L} \in H$ para todo entero positivo L entonces $(a_k)_{k=1}^{\infty} \in H$.

  Cada elemento de H es un *historial*; cada componente de un historial es una *acción* tomada por un jugador. Un historial $(a_k)_{k=1,\dots,K} \in H$ es *terminal* si es finita o si no existe a_{K+1} tal que $(a_k)_{k=1,\dots,K+1} \in H$. El conjunto de acciones disponibles luego del historial no terminal h será notado $A(h) = \{a : (h, a) \in H\}$ y denotaremos T al conjunto de historiales terminales.
- Una *función jugador* $J : H \setminus T \to \mathcal{N}$ que indica el jugador $J(h)$ que debe tomar una acción a continuación para un historial no terminal h.
- Una función de pagos $U_i : T \to \mathbb{R}$ para cada $i \in \mathcal{N}$ que indica la utilidad del jugador i para cada historial terminal. Cada jugador intentará maximizar estos pagos.

A veces en conveniente especificar la estructura de un juego extensivo sin explicitar las preferencias que tienen los jugadores, que formalizamos en las funciones de pago.

Nos referiremos a la tupla $(\mathcal{N}, H, J)$ cuyas componentes satisfacen las primeras 3 condiciones como un *juego en forma extensiva con información perfecta* y a menos que aclaremos lo contrario los llamaremos simplemente "juegos extensivos".

Si el conjunto H es finito entonces el juego se dice *finito*. Si longitud máxima de cualquier historial es finita entonces el juego tiene *horizonte finito*. Si h es una historia de longitud k notamos (h, a) al historial de longitud $k + 1$ que consiste de h seguido de a.

Interpretaremos estos juegos de la siguiente manera: siempre que estemos en un historial h que no sea terminal, el jugador $J(h)$ elige una acción del conjunto

$$A(h) = \{a \ : \ (h, a) \in H\}.$$

El historial vacío, al cual nos referiremos como *historial inicial*, es el punto inicial del juego. En este punto, y siguiendo la regla inicialmente mencionada, el jugador $J(\varnothing)$ elige una acción de $A(\varnothing)$. Una vez que el primer jugador toma alguna acción

a_1 de este conjunto, el jugador $J(a_1)$ elige una acción de $A(a_1)$ y esto determina el siguiente jugador, y el juego continúa de esta manera. Si $A(h) = \varnothing$ entonces h es un historial terminal, y viceversa. Es importante notar que no se puede tomar una acción luego de un historial infinito, y por lo tanto un historial infinito es terminal.

En general, usamos árboles dirigidos para representar la forma extensiva de un juego, asignando un nodo a cada historial, de modo que la raíz sea el comienzo y las hojas los estados terminales. Una acción será una arista entre dos nodos y los historiales se corresponden a caminos entre los vértices del grafo. De hecho esta representación sugiere una definición alternativa con dicho árbol como estructura básica, algo que puede encontrarse en la literatura frecuentemente, como en [8].

Una estrategia para un jugador en un juego extensivo será una política (o plan) en la que esté establecido qué acción tomar para cualquier historial después del cual sea su turno de jugar.

Definición 2.2. Una *estrategia* para el jugador $i \in \mathcal{N}$ en un juego extensivo con información perfecta $(\mathcal{N}, H, J, U_i)$ es una función que asigna una acción a perteneciente a $A(h)$ para cada historial no terminal $h \in H \setminus Z$ tal que $J(h) = i$.

Notemos que la noción de estrategia para un jugador en el juego $(\mathcal{N}, H, J, U_i)$ depende únicamente de la forma del juego $(\mathcal{N}, H, J)$.

Para cada perfil de estrategias $e = (e_i)_{i \in \mathcal{N}}$ del juego $(\mathcal{N}, H, J, U_i)$ definimos el *resultado* $R(e)$ de e como el historial terminal que resulta cuando cada jugador $i \in \mathcal{N}$ hace lo indicado por e_i. Esto es, $R(e)$ es el historial $(a_1, ..., a_K) \in T$ tal que para $0 \leq k \leq K$ se tiene $e_{J(a_1,...,a_k)}(a_1, ..., a_k) = a_{k+1}$.

Del mismo modo que en los juegos en forma estratégica, podemos definir una *estrategia mixta* como una distribución de probabilidad sobre el conjunto de estrategias (puras), aunque esta definición cobra más sentido cuando se estudian situaciones en las que los jugadores no están perfectamente informados.

2.1.1. Equilibrio de Nash

El primer concepto de solución que podemos definir para juegos extensivos no tiene en consideración la estructura secuencial del juego: se considera a las estrategias como opciones que son tomadas al comienzo del juego por única vez y hasta el final.

Definición 2.3. Un *equilibrio de Nash para un juego extensivo con información perfecta* $(\mathcal{N}, H, J, U_i)$ es un perfil de estrategias e^* tal que para todo jugador $i \in \mathcal{N}$ tenemos

$$U_i(R(e^*_{-i}, e^*_i)) \geq U_i(R(e^*_{-i}, e_i))$$

para cualquier estrategia e_i del jugador i.

De manera equivalente, podemos definir un equilibrio de Nash para un juego extensivo Γ, como el equilibrio de Nash correspondiente al juego en forma normal derivado de Γ que se define de la siguiente manera:

Definición 2.4. La forma estratégica (o normal) del un juego extensivo con información perfecta $\Gamma = (\mathcal{N}, H, J, U_i)$ es el juego $(\mathcal{N}, X_i, J_i)$ donde para cada jugador $i \in \mathcal{N}$, X_i es el conjunto de estrategias del jugador i en Γ y el pago de cada perfil de acciones es el correspondiente al pago del historial terminal generado por ese perfil de acciones en el juego extensivo.

De la definición anterior podemos deducir que el conjunto de equilibrios de Nash de un juego extensivo con información perfecta es el conjunto de Equilibrios de Nash de su forma estratégica.

2.1.2. Equilibrios perfectos en subjuegos

La consideración secuencial en la toma de decisiones por parte de los jugadores podría conllevarnos a una situación en la que incluso jugando un equilibrio de Nash definido como recién, alguno de los jugadores encuentre una ganancia al cambiar de acción. Es por eso que definiremos el concepto de equilibrio perfecto en subjuegos. Comencemos por introducir los subjuegos.

Definición 2.5. El subjuego del juego extensivo con información perfecta $\Gamma = (\mathcal{N}, H, J, U_i)$ que sigue el historial h es el juego extensivo $\Gamma(h) = (\mathcal{N}, H|_h, J|_h, U_i|_h)$, donde $H|_h$ es el conjunto de secuencias h' de acciones tales que $(h, h') \in H$, $J|_h$ se define como $J|_h(h') = J(h, h')$ para cada $h' \in H|_h$, y $U_i|_h$ es la misma función de pagos restringida a los historiales que comienzan con h.

Para la definición de equilibrio que introduciremos ahora debemos asumir que los jugadores son racionales, esto es, que la acción a tomar que indica la estrategia de cada jugador debe ser óptima, dadas las estrategias de los demás jugadores, después de cualquier historial. Dada una estrategia e_i para el jugador i y un historial h en el juego extensivo Γ, notaremos $e_i|_h$ a la estrategia que inducida por e_i en el subjuego $\Gamma(h)$, es decir, $e_i|_h(h') = e_i(h, h')$ para cada $h' \in H|_h$. Usaremos R_h para referirnos a la función de resultado de $\Gamma(h)$.

Definición 2.6. Un *equilibrio perfecto en subjuegos* para el juego extensivo con información perfecta $\Gamma = (\mathcal{N}, H, J, U_i)$ es una estrategia e^* tal que para todo jugador $i \in \mathcal{N}$ y todo historial no terminal $h \in H \setminus T$ que cumplen $J(h) = i$ se tiene:

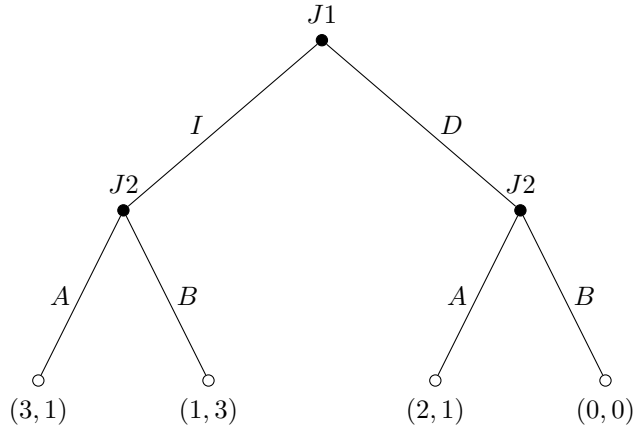
$$U_i(R_h(e^*_{-i}|_h, e^*_i|_h)) \geq U_i(R_h(e^*_{-i}|_h, e_i))$$

para cualquier estrategia e_i del jugador i en el subjuego $\Gamma(h)$.

De manera equivalente, podemos definir un equilibrio perfecto en subjuegos como una estrategia e^* para la cual $e^*|_h$ es un equilibrio de Nash en el subjuego $\Gamma(h)$ para todo historial h.

Para verificar que e^* es una estrategia de equilibrio perfecto en subjuegos, la definición 2.6 nos indica que debemos chequear, para cada jugador i y cada subjuego, que ninguna estrategia conduce a algún pago que el jugador i prefiera.

Ejemplo. Consideremos el siguiente juego extensivo para dos jugadores, $\mathcal{N} = \{1, 2\}$:



Aquí, la raíz del árbol representa el historial vacío y como indica la etiqueta, es el turno del Jugador 1, es decir, $J(\varnothing) = 1$. Este tendrá que decidir entre sus dos posibles acciones: Izquierda (I) y Derecha (D). Esto lo notamos $A(\varnothing) = \{I, D\}$. Los nodos en el segundo nivel representan el momento en que el Jugador 2 debe tomar una decisión, $J(I) = J(D) = 2$, en ambos sus posibles acciones son las mismas, $A(I) = A(D) = \{A, B\}$. Las cuatro hojas del árbol representan los historiales terminales, donde ambos jugadores han tomado ya sus decisiones. Entonces se reparten los pagos según la dupla correspondiente al nodo, siendo la primer coordenada el pago del Jugador 1 y la segunda del Jugador 2. De esta manera, $U_1((I, A)) = 3$, $U_2((I, A)) = 1$, $U_1((I, B)) = 1$, $U_2((I, B)) = 3$, $U_1((D, A)) = 2$, $U_2((D, A)) = 1$, $U_1((D, B)) = 0$, $U_2((D, B)) = 0$, quedando los pagos del juego definidos, y con ellos el juego en forma extensiva con todas sus componentes.

El perfil de estrategias $e = (I, (B, B))$ es decir, el Jugador 1 tomando la acción I y el Jugador 2 eligiendo B independientemente de lo que elija el Jugador 1, constituye un equilibrio de Nash, puesto que ningún Jugador se vería beneficiado cambiando unilateralmente su estrategia. Sin embargo, si el Jugador 1 decidiera jugar D entonces la estrategia del Jugador 2 no sería óptima, pues le convendría

elegir A y obtener una utilidad de 1 en vez de 0. Estas situaciones son las que justifican el uso de los equilibrios perfectos en subjuegos: la estrategia (B, B) para el Jugador 2 no es óptima cuando nos restringimos al subjuego inducido por D, que notamos $\Gamma(D)$. Con esto, el perfil de estrategias $e^* = (I, (B, A))$, además de ser un equilibrio de Nash, es un equilibrio perfecto en subjuegos para nuestro ejemplo.

El siguiente lema nos indica que en los casos en que el juego es de horizonte finito podemos restringirnos a las estrategias alternativas que difieren de e_i^* en las acciones que ésta indica despues de solo un historial. Específicamente, una estrategia es un equilibrio perfecto en subjuegos si y sólo si para cada subjuego el jugador que hace el primer movimiento no puede obtener un mejor pago cambiando sólo su acción inicial.

Notemos con $\ell(\Gamma)$ a la longitud del historial mas corto para Γ, a quien nos referiremos como la lognitud del juego Γ.

Lema 2.7 (La propiedad del desvío único)**.** *Sea $\Gamma = (\mathcal{N}, H, J, U_i)$ un juego extensivo de horizonte finito con información perfecta. La estrategia e^* es un equilibrio perfecto en subjuegos para Γ si y sólo si para todo jugador $i \in \mathcal{N}$ y todo historial $h \in H$ con $J(h) = i$ se tiene que*

$$U_i(R_h(e^*_{-i}|_h, e^*_i|_h)) \geq U_i(R_h(e^*_{-i}|_h, e_i))$$

para toda estrategia e_i del jugador i en el subjuego $\Gamma(h)$ que difiere de $e^*_i|_h$ solo en la acción que está indicada para continuar el juego después del historial inicial de $\Gamma(h)$.

Demostración. Si e^* es un equilibrio perfecto en subjuegos para Γ entonces satisface la condición. Para probar la vuelta, supongamos que e^* no es un equilibrio perfecto en subjuegos y que el jugador i puede cambiar su estrategia unilateralmente para su beneficio en el subjuego $\Gamma(h')$. Esto es que existe una estrategia e_i para el Jugador i en $\Gamma(h')$ para la cual $e_i(h) \neq (e^*_i|_{h'})(h)$ para un número de historiales h menor o igual que la longitud de $\Gamma(h')$. Como el horizonte de Γ es finito entonces este número también. De entre todos los desvíos beneficiosos para i en $\Gamma(h')$ elijamos aquella e_i para la cual la cantidad de historiales h tales que $e_i(h) \neq (e^*_i|_{h'})(h)$ sea mínima. Sea ahora h^* el historial de menor longitud de entre todos los historiales h de $\Gamma(h')$ para el cual se cumple $e_i(h) \neq (e^*_i|_{h'})(h)$. Entonces el historial inicial de $\Gamma(h^*)$ es el único historial en $\Gamma(h^*)$ para el cual la acción indicada por e_i difiere de la indicada por $e^*_i|_{h'}$. Además, $e_i|_{h^*}$ es un desvío beneficioso en $\Gamma(h^*)$ para i, pues de otro modo habría otro desvío en $\Gamma(h')$ que difiere de $e^*_i|_{h'}$ después de un número menor de historiales del que lo hace e_i. Luego, $e_i|_{h^*}$ es un desvío redituable para i en $\Gamma(h^*)$ que difiere de $e^*_i|_{h^*}$ solo en la acción indicada después del historial inicial de $\Gamma(h^*)$. □

Haremos ahora una prueba constructiva del Teorema de Kuhn sobre la existencia de equilibrios perfectos: para cada historial no terminal maximal del juego elegimos una acción óptima para el jugador que tiene el turno de mover y reemplazamos estos historiales con otros terminales cuyos pagos son los resultantes cuando se toma esa elección. Repetimos este procedimiento hacia atrás, hasta el comienzo del juego.

Teorema 2.8. *Todo juego extensivo finito con información perfecta tiene un equilibrio perfecto en subjuegos.*

Demostración. Sea $\Gamma = (\mathcal{N}, H, J, (U_i))$ un juego extensivo finito con información perfecta. Construiremos un equilibrio perfecto en subjuegos para Γ haciendo inducción en $\ell(\Gamma(h))$ y simultáneamente definiremos una función r que asocie a cada $h \in H$ con un historial terminal de manera que este historial sea el equilibrio perfecto en subjuegos resultante para el subjuego $\Gamma(h)$.

Si $\ell(\Gamma(h)) = 0$, es decir que h es terminal, definamos $r(h) = h$. Ahora supongamos que r está definida para todo $h \in H$ con $\ell(\Gamma(h)) \leq k$ y sea h^* un historial para el cual $\ell(\Gamma(h^*)) = k + 1$ y supongamos que $J(h^*) = i$. Como $\ell(\Gamma(h^*)) = k + 1$ entonces $\ell(\Gamma(h^*, a)) \leq k$ para cualquier acción $a \in A(h^*)$. Definamos $e_i(h^*)$ de modo que maximice $U_i(r(h^*, a))$ sobre las acciones $a \in A(h^*)$ y definamos $r(h^*) = r(h^*, e_i(h^*))$.

De esta manera, hemos definido inductivamente un perfil de estrategias e para Γ que cumple que para cada subjuego el jugador que hace el primer movimiento lo hace óptimamente, entonces por la propiedad del desvío único (Lema 2.7) e resulta un equilibrio perfecto en subjuegos para Γ. □

El procedimiento que hicimos en la demostración se conoce como inducción hacia atrás y presenta el procedimiento habitual que se utiliza para resolver juegos extensivos finitos cuando la información es perfecta. El inconveniente se da cuando la longitud de los subjuegos excede la capacidad de cómputo, como por ejemplo en el ajedrez, o cuando la información es imperfecta o incompleta porque esto implica tomar decisiones a través de nodos en los que la información necesaria no está disponible [21].

2.2. Juegos de información imperfecta y el enfoque Bayesiano

En esta sección trataremos juegos extensivos en los que los jugadores no cuentan con toda la información que podrían desear tener, es decir, juegos de información

imperfecta, para los cuales extenderemos la noción de equilibrio perfecto en subjuegos. En ese punto nos encontraremos con el concepto de equilibrio secuencial, que analizaremos posteriormente en el caso particular de los juegos bayesianos.

La siguiente definición generaliza la de un juego extensivo con información perfecta (2.1) permitiendo a los jugadores estar desinformados sobre eventos pasados a la hora de tomar una decisión. Esto nos permite también agregar a nuestro modelo la posibilidad de que el azar influya en las acciones de los jugadores o en la distribución de distintos atributos que hagan a la estructura del juego, introduciendo incertezas exógenas que habrá que analizar a la hora de encontrar estrategias óptimas.

Definición 2.9. Un *Juego extensivo (con información imperfecta)* consta de las siguientes componentes:

- Un conjunto $\mathcal{N}$ (el conjunto de jugadores).
- Un conjunto de secuencias H (finito o no) que satisface:
  - La secuencia vacía $\varnothing$ pertenece a H.
  - Si $(a_k)_{k=1,\dots,K} \in H$ (donde K podría ser infinito) y $L \leq K$ entonces $(a_k)_{k=1,\dots,L} \in H$.
  - Si $(a_k)_{k=1}^{\infty}$ cumple que $(a_k)_{k=1,\dots,L} \in H$ para todo entero positivo L entonces $(a_k)_{k=1}^{\infty} \in H$.

  Cada elemento de H es un *historial*; cada componente de un historial es una *acción* tomada por un jugador. Un historial $(a_k)_{k=1,\dots,K} \in H$ es *terminal* si es infinito o si no existe a_{K+1} tal que $(a_k)_{k=1,\dots,K+1} \in H$. Denotaremos T al conjunto de historiales terminales. El conjunto de acciones disponibles luego del historial no terminal h será notado $A(h) = \{a : (h, a) \in H\}$.
- Una *función jugador* $J : H \setminus T \to \mathcal{N}$ que indica el jugador $J(h)$ que debe tomar una acción a continuación dado que se sucedió el historial no terminal h.
- Una función f_c que se asocia cada historial h tal que $J(h) = c$ con una medida de probabilidad $f_c(\cdot|h)$ sobre $A(h)$, donde cada una de esas medidas es independiente de las demás. ($f_c(a|h)$ será la probabilidad de que ocurra a luego del historial h.)
- Para cada jugador $i \in \mathcal{N}$ una partición $\mathcal{I}_i$ del conjunto $\{h \in H : J(h) = i\}$ con la propiedad de que $A(h) = A(h')$ siempre que h y h' forman parte del mismo miembro de la partición. Para cada $I_i \in \mathcal{I}_i$ notaremos con $A(I_i)$

al conjunto $A(h)$ y con $J(I_i)$ al jugador $J(h)$ para cualquier $h \in I_i$. ($\mathcal{I}_i$ es la *partición de la información* correspondiente al jugador i, y un conjunto $I_i \in \mathcal{I}_i$ es un conjunto de información del jugador i)

- Una función de pagos $U_i : T \to \mathbb{R}$ para cada $i \in \mathcal{N}$ que indica la utilidad del jugador i para cada historial terminal.

Nos referiremos a la tupla $(\mathcal{N}, H, J, f_c, (\mathcal{I}_i)_{i\in\mathcal{N}})$ (excluyendo los pagos del juego) cuyas componentes satisfacen estas condiciones como un *juego en forma extensiva* (o la forma extensiva de un juego).

Interpretaremos los historiales de cualquier miembro de $\mathcal{I}_i$ como indistinguibles para el jugador i. De esta manera, el juego modela la situación en la que luego de cada historial $h \in I_i \in \mathcal{I}_i$ el jugador i es conciente de que ocurrió algún historial en I_i pero no sabe que éste es precisamente el historial h. La condición de que $A(h) = A(h')$ cuando h y h' están en el mismo miembro de $\mathcal{I}_i$ captura la idea de que si fueran distintos entonces el jugador i podría deducir, cuando se encuentra con $A(h)$, que el historial no fué h', contradiciendo nuestra interpretación de $\mathcal{I}_i$. Las particiones del conjunto de información de los jugadores son una pata importante del modelo: un jugador distingue historiales en diferentes elementos de su partición sin que las acciones que observa le produzcan distorsiones. Mientras transcurre el juego, cada participante puede hacer conjeturas a partir del comportamiento de sus oponentes, y de esta manera inferir información que refine su conjunto.

Definición 2.10. Una *estrategia pura* del jugador $i \in \mathcal{N}$ en un juego extensivo $(\mathcal{N}, H, J, f_c, (\mathcal{I}_i), (U_i))$ es una funcion que asigna una acción de $A(I_i)$ a cada conjunto de información $I_i \in \mathcal{I}_i$.

Como en los juegos extensivos con información perfecta, podemos asociar a cada juego extensivo un juego en forma estratégica, ver 2.4.

La forma en que modelamos un juego extensivo nos permite considerar una gran variedad de situaciones y entornos de información. En particular, podríamos capturar la situación (que no estudiaremos en esta tesis) en la que un jugador pierda información que tenía antes, como si olvidase lo que jugó en el pasado. Es por eso que nos referiremos a juegos en los que en cualquier momento cada jugador recuerda todo lo que sabía como *juegos con memoria perfecta*.

2.2.1. Estrategias mixtas y conductuales

En 2.10 definimos estrategias puras para un juego extensivo. Hay dos maneras de modelar la posibilidad de que las acciones de un jugador dependan de factores

aleatorios:

Definición 2.11. Una *estrategia mixta para el jugador i* en el juego extensivo $\Gamma = (\mathcal{N}, H, J, f_c, (\mathcal{I}_i), (U_i))$ es una medida de probabilidad sobre el conjunto de estrategias puras del jugador i. Una *estrategia conductual para el jugador i* es una colección $(\beta_i(I_i))_{I_i \in \mathcal{I}_i}$ de medidas de probabilidad independientes, donde $\beta_i(I_i)$ tiene soporte en $A(I_i)$.

Dados un historial $h \in I_i \in \mathcal{I}_i$ y una acción $a \in A(h)$ notaremos $\beta_i(h)(a)$ a la probabilidad $\beta_i(I_i)(a)$ asignada por $\beta_i(I_i)$ a la acción a.

De esta manera, la definición de estrategia mixta es equivalente a la que definimos anteriormente para juegos estratégicos. Sin embargo, una estrategia conductual especifica una medida de probabilidad sobre las acciones disponibles para el jugador i en cada uno de sus conjuntos de información. Así, el jugador puede variar su comportamiento de dos formas: jugando aleatoriamente alguna de sus estrategias puras, o planeando de manera mas general una decisión estocástica en cada uno de los puntos en los que tenga que tomar una acción.

Para cualquier perfil $e = (e_i)_{i \in \mathcal{N}}$ de estrategias mixtas o conductuales de un juego extensivo, podemos definir el *resultado $R(e)$ de e* como la distribución de probabilidades sobre los historiales terminales que resulta cuando cada jugador i sigue las instrucciones estipuladas por e_i. Para un juego finito podemos formular una definición precisa como sigue.

Dado un historial $h = (a_1, ..., a_k)$ definamos una estrategia e_i para el jugador i como *consistente* con h si para cada subhistorial $(a_1, ..., a_l)$ de h tal que $J(a_1, ..., a_l) = i$ se tiene que $e_i(a_1, ..., a_l) = a_{l+1}$. Sea $\pi_i(h)$ la suma de las probabilidades correspondientes a e_i de todas las estrategias puras para el jugador i que son consistentes con h. Si h es un historial en el que el jugador i no toma ninguna acción entonces $\pi_i(h) = 1$. Luego, para cada perfil e de estrategias mixtas la probabiliadad que $R(e)$ asigna a cada historial no terminal h es $\prod_{i \in \mathcal{N} \cup \{c\}} \pi_i(h)$. Si β es un perfil de estrategias conductuales entonces la probabilidad que $R(\beta)$ asigna al historial terminal $h = (a_1, ..., a_K)$ es $\prod_{k=0}^{K-1} \beta_{J(a_1,...,a_k)}(a_1, ..., a_k)(a_{k+1})$ donde para $k = 0$ el historial $(a_1, ..., a_k)$ es el historial inicial.

Dos estrategias (mixtas o conductuales) se dicen *equivalentes en resultado* si para cualquier colección de estrategias puras de los demás jugadores, ambas inducen el mismo resultado. Entonces:

Proposición 2.12. *Para cualquier estrategia mixta de un juego extensivo finito con memoria perfecta existe una estrategia conductual equivalente en resultado.*

Demostración. Sea e_i una estrategia mixta del Jugador i. Como recién, para cada historial h sea $\pi_i(h)$ la suma de las probabilidades establecidas por e_i de todas las

estrategias puras para el jugador i que son consistentes con h. Sean h y h' dos historiales en el mismo conjunto de información I_i del Jugador i, y sea $a \in A(h)$. Como el juego tiene memoria perfecta, los conjuntos de acciones de i en h y h' son los mismos y por lo tanto $\pi_i(h) = \pi_i(h')$. Más aún, $\pi_i(h, a) = \pi_i(h', a)$, pues para cualquier estrategia pura del Jugador i la acción a es tomada después de que transcurrió h si y sólo si es tomada después de h'. Con esto, podemos definir una estrategia conductual β_i para el Jugador i mediante $\beta_i(I_i)(a) = \pi_i(h, a)/\pi_i(h)$ para todo $h \in I_i$ que cumpla $\pi_i(h) > 0$ y así $\sum_{a \in A(h)} \beta_i(I_i)(a) = 1$. La definición de $\beta_i(I_i)(a)$ cuando $\pi_i(h) = 0$ es irrelevante.

Veamos que así definida, β_i es equivalente en resultado a e_i. Sean η_{-i} una colección de estrategias puras para los demás jugadores y $h \in T$ un historial terminal. Si h incluye movimientos inconsistentes con η_{-i} entonces la probabilidad de h es cero tanto para e_i como para β_i. Supongamos entonces que todos los movimientos de los jugadores (distintos de i) en h son consistentes con η_{-i}. Si h incluye una acción después de algún subhistorial $h' \in I_i$ de h que es inconsistente con e_i entonces $\beta_i(I_i)$ asigna probabilidad positiva a esa acción, y por lo tanto la probabilidad para β_i de h es cero. Por último, si h es consistente con e_i entonces $\pi_i(h') > 0$ para cualquier subistorial h' de h y la probailidad que tiene h de acuerdo con β_i es el producto de todas las $\pi_i(h', a)/\pi_i(h')$ sobre todos los (h', a) que son subistoriales de h, y este producto es exactamente $\pi_i(h)$, la probabilidad de tiene h según e_i. □

2.2.2. Estrategias y Creencias

Recordemos que un equilibrio perfecto en subjuegos para un juego extensivo con información perfecta es un perfil de estrategias en el que la estrategia de cada jugador es óptima (dadas las estrategias de los demás jugadores) para cualquier historial en el que dicho jugador sea el próximo que debe tomar una acción, ocurra o no dicho historial mientras que los jugadores sigan la receta indicada por dicho perfil. Podemos aplicar naturalmente esta idea en juegos extensivos con información imperfecta buscando la estrategia óptima dentro de cada uno de los conjuntos de información de los jugadores.

A diferencia de las soluciones que estudiamos anteriormente, que eran equilibrios formados únicamente por el perfil de estrategias, nos enfocaremos en un concepto de solución con dos componentes: los equilibrios secuenciales que introduciremos consistirán de un perfil de estrategias y de un *sistema de creencias*, que especificará, para cada conjunto de información, las suposiciones que hacen los jugadores que tienen que tomar una acción sobre ese conjunto de información acerca del historial que ocurrió. Uno de estos pares será llamado una *evaluación*, y

el sistema de creencias será una colección de medidas de probabiliadad, una para cada conjunto de información. Notemos que la noción de evaluación coincide con la de perfil de estrategias para un juego con información perfecta, dado que en un juego de ese tipo los conjuntos de información constan de un único elemento y por lo tanto hay un único sistema de creencias posible.

Para definir los equilibrios perfectos en subjuegos pedimos que la estrategia de cada jugador fuera óptima para el subjuego generado por cualquier historial, la hipótesis que tendremos que considerar ahora es la de *racionalidad secuencial*: para cada conjunto de información de cada jugador i, su estrategia (conductual) indicará una respuesta óptima a las estrategias de los demás jugadores, dadas las creencias del jugador i sobre ese conjunto de información.

Hasta ahora no hemos impuesto ninguna restricción sobre las creencias de los jugadores. Incluiremos aquí las consideradas en [20] aunque es posible encontrar gran variedad en la literatura.

- *Consistencia con las estrategias* Debemos pedir que el sistema de creencias sea consistente con el perfil de estrategias en el sentido de que para cualquier conjunto de información consistente con las estrategias de los jugadores, la creencia acerca del historial que ocurrió debe poder deducirse de las estrategias usando reglas bayesianas.

- *Consistencia estructural* Incluso aunque un conjunto de información no sea alcanzado, si todos los jugadores adhieren a sus estrategias, requeriremos que las creencias de los jugadores se deduzcan de *algún* perfil de estrategias (alternativo) utilizando reglas bayesianas. (Nos referimos a esta condición como estructural pues no depende de los pagos de los jugadores ni de la estrategia de equilibrio)

- *Creencias comunes* En general, cuando nos referimos a soluciones teóricas de juegos debemos incluir cualquier asimetría en la descripción del juego: asumimos que los jugadores analizan la situación del mismo modo. En el contexto del equilibrio perfecto en subjuegos esto nos conduce a requerir que las creencias de los jugadores sobre los planes del jugador i en caso de que ocurra un evento inesperado sean las mismas. En el contexto en que nos encontramos ahora esto se traduce en asumir que todos los jugadores comparten la misma creencia acerca de la causa de cualquier evento inesperado.

2.2.3. Equilibrio secuencial

Definición 2.13. Una *evaluación* en un juego extensivo es un par (β, μ), con β un perfil de estrategias conductuales y μ una función que asigna a cada conjunto de

información una medida de probabilidad sobre el conjunto de historiales en dicho conjunto de información.

Sea (β, μ) una evaluación en $\Gamma = (\mathcal{N}, H, J, f_c, (\mathcal{I}_i), (U_i))$. Llamaremos a μ un *sistema de creencias* e interpretaremos que $\mu(I)(h)$ es la probabilidad que el jugador $J(I)$ asigna al historial $h \in I$, condicional a que ocurrió I.

Diremos que una evaluación es *secuencialmente racional* si para cada $i \in \mathcal{N}$ y para cada conjunto de información $I_i \in \mathcal{I}_i$ la estrategia del jugador i es la mejor respuesta a las estrategias de los otros jugadores dada la creencia de i en I_i. Para formalizar esto último definamos el *resultado* $R(\beta, \mu|I)$ *de* (β, μ) *condicional a* I como la distribución sobre los historiales terminales determinada por β y μ condicional a que ocurra I, de la siguiente manera: Sea $h^* = (a_1, ..., a_K)$ un historial terminal y sea $h = (a_1, ..., a_L) \in I$ para algún $L < K$. Entonces:

- $R(\beta, \mu|I)(h^*) = 0$ si no hay ningún subhistorial de h^* en I.
- $R(\beta, \mu|I)(h^*) = \mu(I)(h) \cdot \prod_{k=L}^{K-1} \beta_{J(a_1,...,a_k)}(a_1, ..., a_k)(a_{k+1})$.

Si I es el conjunto de información consistente del historial inicial entonces $R(\beta, \mu|I)$ es simplemente el resultado $R(\beta)$ definido en la sección anterior.

Definición 2.14. Sea $\Gamma = (\mathcal{N}, H, J, f_c, (\mathcal{I}_i), (U_i))$ un juego extensivo con memoria perfecta. Diremos que la evaluación (β, μ) es *secuencialmente racional* si para cada jugador $i \in \mathcal{N}$ y cada conjunto de información $I_i \in \mathcal{I}_i$ se tene que

$$U(R(\beta, \mu|I)) \geq U(R(\beta_{-i}, \beta_i'), \mu|I_i))$$

para toda estrategia β_i' del jugador i.

Definición 2.15. Sea $\Gamma = (\mathcal{N}, H, J, f_c, (\mathcal{I}_i), (U_i))$ un juego extensivo con memoria perfecta. Una evaluación (β, μ) se dice *consistente* si existe una sucesión $((\beta_n, \mu_n))_{n=1}^{\infty}$ de evaluaciones que convergen a (β, μ) en el espacio euclideo y con las propiedades de que cada perfil de estrategias β^n es completamente mixto y cada sistema de creencias μ_n se deduce de β^n usando reglas de Bayes.

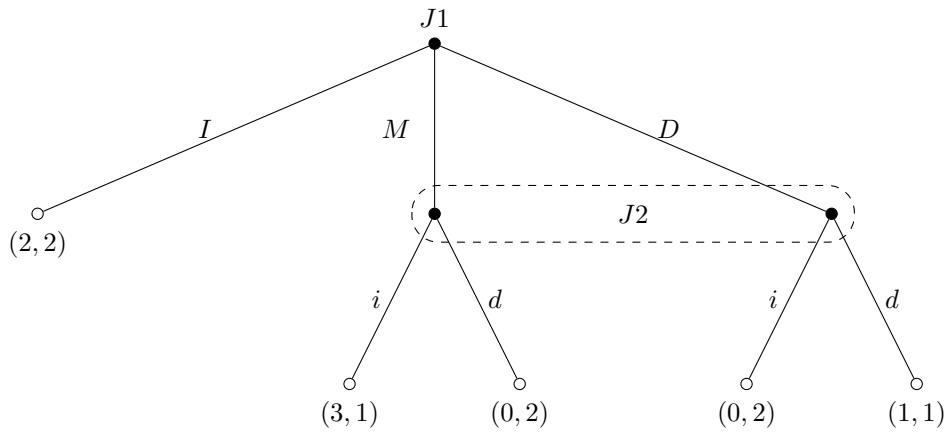
El fundamento principal para considerar la consistencia como hipótesis es que la probabilidad condicionada a eventos de probabilidad cero debe aproximar a las probabilidades que se derivan de estrategias completamente mixtas, es decir, que asignan probabilidad positiva a todas las acciones. En [16] podemos encontrar los tecnicismos detrás de la definición.

Definición 2.16. Una evaluación será un *equilibrio secuencial* de un juego extensivo finito con memoria perfecta si es secuencialmente racional y consistente.

Observación 6. Si (β, μ) es un equilibrio secuencial entonces β es un equilibrio de Nash. Más aún, en un juego extensivo con información perfecta (β, μ) es un equilibrio secuencial si y sólo si β es un equilibrio perfecto en subjuegos.

Definición 2.17. Un sistema de creencias μ para un juego extensivo con memoria perfecta se dice *estructuralmente consistente* si para cada conjunto de información I existe un perfil de estrategias β tal que I tiene probabiliad positiva de ser alcanzado para β y $\mu(I)$ se deduce de β vía inferencia bayesiana.

Ejemplo. Consideremos el siguiente juego extensivo con información imperfecta:



Analicemos los equilibrios de este ejemplo. El conjunto de información del Jugador 1 es simplemente el nodo raíz, y el del Jugador 2 es el conjunto $I_2 = \{M, D\}$. Es sobre estos conjuntos que debemos indicar las creencias de los Jugadores, aunque en este ejemplo esto resulta no trivial sólo para I_2.

Es fácil ver que (I, d) es un Equilibrio de Nash. Pero sucede que si el Jugador 1 elige I entonces el conjunto de información del Jugador 2 no es alcanzado. Es de esta manera que notamos la importancia de las estrategias conductuales, en las que debemos indicar la acción óptima incluso cuando el conjunto de información no sea alcanzado si el Jugador 1 sigue lo indicado. Para esto debemos analizar el conjunto de creencias (del Jugador 2) sobre cómo fué alcanzado este conjunto.

Así, la acción d es óptima si le asignamos probabilidad de al menos $1/2$ a M, y si la probabilidad $\mu(I_2)(M) < 1/2$ entonces la acción óptima del Jugador 2 será i.

La evaluación (β, μ) en la que $\beta_1(I) = 1$, $\beta_2(d) = 1$, y $\mu(I_2)(M) = \alpha$ para cualquier $\alpha \in [0, 1]$ es consistente por ser el límite con ε tendiendo a cero de la sucesión de evaluaciones $(\beta^\varepsilon, \mu^\varepsilon)$, donde $\beta_1^\varepsilon = (1 - \varepsilon, \alpha\varepsilon, (1 - \alpha)\varepsilon)$, $\beta_2^\varepsilon = (\varepsilon, 1 - \varepsilon)$ y $\mu^\varepsilon(I_2)(M) = \alpha$ para todo ε. Si $\alpha \geq 1/2$ entonces la evaluación es además secuencialmente racional, por lo tanto se trata de un equilibrio secuencial.

2.2.4. Juegos Bayesianos

Concluiremos analizando los juegos extensivos Bayasianos, donde podremos definir un equilibrio muy similar al secuencial aunque un poco más simple. En este tipo de juegos forma parte del conocimiento de los jugadores la secuencia de acciones que fueron tomadas; la única incerteza se representa en un movimiento inicial en el que los jugadores reciben una distribución de probabilidades sobre la información relevante de sus pagos personales. Esto además se considera de manera tal que la información que recibe cada jugador no revela ningún otro tipo de información acerca de ningún otro jugador. Diremos que este movimiento inicial asigna un *tipo* a cada jugador, y diremos que el jugador i que recibe la información θ_i es de *tipo* θ_i.

Definición 2.18. Un juego extensivo *Bayesiano* está compuesto por una tupla $(\Gamma, (\Theta_i), (\pi_i), (U_i))$ donde:

- $\Gamma = (\mathcal{N}, H, J)$ es un juego extensivo con información perfecta y para cada jugador $i \in \mathcal{N}$:
- Θ_i es un conjunto finito (el conjunto de todos los posibles *tipos* del jugador i) y notamos $\Theta = \prod_{i \in \mathcal{N}} \Theta_i$
- π_i es una medida de probabilidad sobre Θ_i para la cual $\pi_i(\theta_i) > 0$ para todo $\theta_i \in \Theta_i$, y las medidas π_i son estocásticamente independientes ($\pi_i(\theta_i)$ representa la probabilidad de que el jugador i sea seleccionado para ser de tipo θ_i)
- $U_i : \Theta \times T \to \mathbb{R}$ es una función de utilidad de von Neumann-Morgenstern [28] ($U_i(\theta, h)$ será el pago para el jugador i cuando es de tipo θ y la historia terminal de Γ es h).

Este modelo representa una situación en la que "la naturaleza", mediante el azar, decide el tipo de cada jugador, pero estos son plenamente concientes, durante el transcurso del juego, de todos los movimientos y decisiones tomadas anteriormente. Podemos asociar a cualquiera de estos juegos un juego extensivo en el que el conjunto de historiales es $\{\varnothing\} \cup (\Theta \times H)$ y cada conjunto de información $I(\theta_i, h) = \{((\theta_i, \theta'_{-i}), h) : \theta'_{-i} \in \Theta_{-i}\}$ para $i \in J(h)$ y $\theta_i \in \Theta_i$ (es decir que la cantidad de historias en el conjunto de información se corresponde con la cantidad de miembros en Θ_{-i}).

Basándonos en los conceptos de equilibrio que definimos en ésta sección, nuestro candidato a solución óptima sería el par

$$((e_i), (\mu_i)) = ((e_i(\theta_i))_{i \in \mathcal{N}, \theta_i \in \Theta_i}, (\mu_i(h))_{i \in \mathcal{N}, h \in H \setminus T})$$

, donde cada $e_i(\theta_i)$ es una estrategia conductual para el jugador i en Γ (la estrategia usada por el jugador i cuando es de tipo θ_i) y cada $\mu_i(h)$ es una medida de probabilidad en Ω_i (esto sería la creencia común, después del historial h, de todos los jugadores que son del mismo tipo que el jugador i). Este par está muy relacionado con lo que definimos anteriormente como una evaluación. El perfil (e_i) presenta una reformulación equivalente de la información como un perfil de estrategias conductuales en el juego extensivo asociado, y el perfil (μ_i) resume las creencias de los jugadores y está adaptada a la suposición de que cada jugador está perfectamente informado acerca de los movimientos anteriores de los demás jugadores y la única incerteza que puede tener es sobre el tipo de los demás jugadores.

Sea e un perfil de estrategias conductuales para Γ. Sea $R_h(e)$ la medida sobre el conjunto de historiales terminales de Γ generadas por e dado que ocurrió el historial h. Definamos $R(e_{-i}, e_i, \mu_{-i}|h)$ como la medida de probabilidad sobre el conjunto de historiales terminales dado que el jugador i usa la estrategia e_i, cada jugador j de tipo θ_j usa la estrategia $e_j(\theta_j)$, el juego se encuentra en el historial h y la probabilidad que i le asigna a θ_{-i} se deduce de $\mu_{-i}(h)$. Esto es, $R(e_{-i}, e_i, \mu_{-i}|h)$ es la componente de la lotería en la que la probabilidad de obtener la lotería $R_h((e_j(\theta_j))_{j\in\mathcal{N}\setminus\{i\}}, s_i)$ es $\prod_{j\in\mathcal{N}\setminus\{i\}} \mu_i(h)(\theta_j)$ para cada $\theta_{-i} \in \Theta_{-i}$.

Definición 2.19. Sea $(\Gamma, (\Theta_i), (\pi_i), (U_i))$ un juego Bayesiano extensivo, con $\Gamma = (\mathcal{N}, H, J)$. El par $((e_i), (\mu_i)) = ((e_i(\theta_i))_{i\in\mathcal{N},\theta_i\in\Theta_i}, (\mu_i(h))_{i\in\mathcal{N},h\in H\setminus T})$, donde $e_i(\theta_i)$ es una estrategia conductual para el jugador i en Γ y $\mu_i(h)$ es una medida de probabilidad sobre el conjunto de tipos Θ_i, es un *equilibrio Bayesiano perfecto* del juego si cumple con las siguientes condiciones:

- *Racionalidad secuencial* Para todo historial no terminal $h \in H\setminus T$, todo jugador $i \in J(h)$, y todo $\theta_i \in \Theta_i$ la medida de probabilidad $R(e_{-i}, e_i(\theta_i), \mu_{-i}|h)$ es tan buena para el tipo θ_i como lo es $R(e_i, e_i, \mu_{-i}|h)$ sea cual sea la estrategia e_i del jugador i. Esto es que la estrategia $e_i(\theta_i)$ para cada tipo y jugador sea óptima despues de cualquier secuencia de eventos.

- *Creencias iniciales bien definidas* $\mu_i(\varnothing) = \pi_i$ para todo $i \in \mathcal{N}$. Es decir que las creencias iniciales de los otros jugadores acerca del tipo de cada jugador i venga dada por la lotería inicial.

- *Creencias determinadas por las acciones* Si $i \notin J(h)$ y $a \in A(h)$ entonces $\mu_i(h, a) = \mu_i(h)$; si $i \in J(h)$, $a, a' \in A(h)$, y $a_i = a'_i$ entonces $\mu_i(h, a) = \mu_i(h, a')$. Esta condición indica que si el jugador i no tiene que mover entonces la acción que se tome no afecta las creencias de los demás jugadores sobre su tipo, y que si i es uno de los que mueve a continuación enonces las creencias de los demás jugadores acerca de su tipo se basan únicamente en lo que

éste haga y no de las acciones de los demás jugadores, algo coherente con la independecia entre las estrategias de los jugadores.

- *Actualización bayesiana* Si $i \in J(h)$ y a_i pertenece al soporte de $e_i(\theta_i)(h)$ para algún θ_i en el soporte de $\mu_i(h)$ entonces para cada $\theta_i' \in \Theta_i$ se tiene que $\mu_i(h,a)(\theta_i')$ se deduce aplicando Bayes, es decir:

$$\mu_i(h,a)(\theta_i') = \frac{e_i(\theta_i')(h)(a_i) \cdot \mu_i(h)(\theta_i')}{\sum_{\theta_i \in \Theta_i} e_i(\theta_i)(h)(a_i) \cdot \mu_i(h)(\theta_i)}$$

  Esto podría contradecir las creencias anteriores de los demás jugadores acerca del tipo de i, implicando una actualización en sus conjeturas: ésta es la base de una futura actualización Bayesiana en el transcurso del juego si ocurre otro conflicto con la estrategia e.

La siguiente proposición indica que un equilibrio secuencial para el juego extensivo asociado al juego Bayesiano finito es equivalente a un equilibrio Bayesiano perfecto en el juego extensivo bayesiano, en el sentido de que induce las mismas creencias y conductas. Una demostración detallada puede encontrarse en [10], donde se explicita cómo construir las creencias comunes inciales a partir del sistema de creencias del equilibrio secuencial.

Proposición 2.20. *Sea (β, μ) un equilibrio secuencial del juego extensivo asiociado al juego finito Bayesiano $((\mathcal{N}, H, J), (\Theta_i), (\pi_i), (U_i))$. Para cada $h \in H$, $i \in J(h)$ y $\theta_i \in \Theta_i$ sea $e_i(\theta_i)(h) = \beta_i(I(\theta_i, h))$. Entonces existe una familia $(\mu_i(h))_{i \in \mathcal{N}, h \in H}$ de medidas de probabilidad sobre Θ_i tal que*

$$\mu(I(\theta_i, h))(\theta, h) = \prod_{j \in \mathcal{N} \setminus \{i\}} \mu_j(h)(\theta_j) \quad \forall \theta \in \Theta,\ h \in H$$

y tal que $((e_i), (\mu_i))$ es un equilibrio Bayesiano perfecto del juego extensivo Bayesiano.

Capítulo 3

Las ventajas de ser el líder

3.1. ¿Extensivo o estratégico?

Recordemos que nuestro objetivo es resolver juegos Stackelberg, es decir, una situación de liderazgo en la que un jugador va a anunciar su estrategia, informando al seguidor antes que elija su acción óptima. Podemos pensar que aquí el seguidor se encuentra con conjuntos de información consistentes de un único elemento, plenamente consciente de la posición del juego extensivo en la que se encuentra. Aún así, el juego puede presentar incertidumbres para el líder sobre sus posibles adversarios, el caso de los juegos Bayesianos. Al mismo tiempo, en el primer capítulo obtuvimos un algoritmo que computaba una estrategia óptima para el líder siempre que esta existiera, o levemente subóptima cuando no.

En principio, el marco teórico de la forma normal es suficiente para definir soluciones que satisfagan nuestros requerimientos finales para juegos de información perfecta, pero introdujimos la forma extensiva por tres razones. Primero, es importante identificar las estructuras sobre las que se definen los subjuegos de la forma extensiva, que no pueden visualizarse en la forma normal. Segundo, nos interesa teóricamente el equilibrio perfecto en subjuegos, un concepto que no se puede definir satisfactoriamente en el marco estructural simple de la forma estratégica. Y tercero, es sobre ésta estructura secuencial sobre la que se desarrolla exitosamente la teoría de juegos con información imperfecta en general, y en particular el enfoque bayesiano y su respectiva solución.

En la sección 1.4 definimos el concepto de equilibrio Stackelberg en estrategias mixtas (Definición 1.17) y lo hicimos tomando una actitud pesimista, es decir, asumimos que en caso de que el seguidor fuera indiferente entre dos estrategias optaría por la que significara menor utilidad para el líder. Esta definición es cohe-

rente con la actitud conservadora que podemos presuponer que toma en la realidad la empresa que actúa en el rol de líder para un modelo económico o una fuerza de seguridad en el caso que pretendemos modelar. Aún así, nos encontramos con el gran inconveniente teórico de la posible no existencia de una solución óptima en estrategias mixtas para el líder, como ejemplificamos al final de la sección mencionada, dado que la utilidad del líder como función del compromiso que toma a jugar cierta estrategia resulta en general discontinua. Mostraremos en el teorema 7 que el pago resultante para el líder se corresponde con un equilibrio perfecto en subjuegos del juego secuencial que definimos en el segundo capítulo. La única diferencia es que los seguidores no actúan (ni pueden hacerlo) de acuerdo a esta actitud pesimista, en cuyo caso el líder podría anunciar una estrategia cercana, levemente subóptima.

Observación 7. Dado un juego extensivo con información perfecta, algunos (pero no todos) los equilibrios perfectos en subjuegos del Juego jerárquico (Stackelberg) corresponden a soluciones del problema de maximización:

$$\max_{(p,q)\in\mathcal{P}\times\mathcal{Q}} U_1(p,q) \text{ sujeto a que } q \in \arg\max_{q'\in\mathcal{Q}} U_2(p,q').$$

Si además, los conjuntos de estrategias $\mathcal{P}$ y $\mathcal{Q}$ son compactos y la funciones de pagos U_1 y U_2 son continuas, este problema tiene solución, como argumentamos en el primer capítulo. [20]

Estamos diciendo que, si asumimos que el seguidor optará por la estrategia que beneficie al líder, entonces siempre existirá una estrategia mixta óptima para el líder, y ésta se corresponderá con un equilibrio perfecto en subjuegos para la representación del juego en forma extensiva. En juegos genéricos de dos jugadores, el pago para el líder es único y es como mínimo el pago del equilibrio de Nash. Esto se debe al hecho de que las respuestas óptimas del seguidor son *únicas en casi todo punto*, relativamente al conjunto de todas las estrategias mixtas del líder. En la próxima sección veremos que es posible que el líder induzca al seguidor a jugar la estrategia que lo beneficie tomando una estrategia arbitrariamente cercana al equilibrio.

Dicho esto, podremos referirnos a *la estrategia óptima del líder* y al equilibrio de Stackelberg fuerte (como lo llama en su trabajo [4]) , que dependiendo de la situación en la que nos encontremos se corresponderá con alguno de los equilibrios perfectos en subjuegos (o equilibrios perfectos Bayesianos dependiendo del contexto).

Los problemas que intentaremos modelar y resolver en ésta tesis tienen poca dificultad teórica, serán juegos finitos (a veces bayesianos) de dos jugadores, y

consideraremos una distribución inicial sobre los distintos tipos de seguidores para indicar distintas matrices de pagos, buscando la estrategia mixta óptima.

3.2. Estrategias inducibles

Es interesante notar que, en el juego jerárquico, el líder puede hacer uso de su rol en su beneficio: "commitment power", el poder de anunciar un compromiso a jugar una estrategia. Las ventajas frente a jugar el juego simultáneo, decidiendo al mismo tiempo que su adversario, son el resultado planteado en [30] al que llegaremos en esta sección: lo peor que puede pasar es que estemos jugando un equilibrio de Nash.

Consideremos un juego bimatricial con matrices de pago A y B para el jugador 1 y el 2 respectivamente, que intentaran maximizar. Para referirnos a ellas de manera más cómoda numeraremos las estrategias puras de ambos jugadores, $X = \{1, ..., x, ..., n\}, Y = \{1, ..., y, ..., m\}$.

De esta manera, cuando hablemos de la columna de A (o de B) correspondiente a la estrategia y del seguidor escribiremos Ay (o By) y pensaremos que y es el vector canónico e_y. Los vectores de estrategias mixtas serán p y q como antes, e indicaremos p_x a la probabilidad de que p le asigna a x, análogamente q_y.

Definición 3.1. Una estrategia (pura) $y \in Y$ de $\mathbf{J}2$ se dice *inducible* si es la única respuesta óptima para alguna estrategia $p \in \mathcal{P}$ de $\mathbf{J}1$, es decir, $R_2(p) = \{y\}$.

Para cada $y \in Y$ consideremos el conjunto de respuesta óptima $R_1(y) \subset \mathcal{P}$ que consta de aquellas estrategias p del Jugador 1 para las cuales y es la respuesta óptima del Jugador 2. Denotemos $a(y)$ el pago maximal al que puede aspirar $\mathbf{J}1$ cuando elige alguna p de $R_1(y)$:

$$a(y) = \text{máx}\{pAy : p \in R_1(y)\}. \tag{3.1}$$

El siguiente lema nos dice que si la estrategia es inducible entonces cualquier pago para el líder es al menos este valor en un equilibrio.

Lema 3.2. *Si y es una estrategia inducible, entonces el pago del líder en cualquier equilibrio Stackelberg será al menos $a(y)$ como lo definimos en (3.1).*

Demostración. Consideremos un par de estrategias (x, y) que constituyan un equilibrio Stackleberg. Supongamos que el máximo $a(y)$ de 3.1 se alcanza en $x \in R_1(y)$, es decir, $a(y) = xAy$. Por otro lado, sea $x' \in R_1(y)$ una estrategia para la cual y es la única respuesta óptima de $\mathbf{J}2$ que existe por ser y inducible. Tenemos que

$x'By > x'By'$ para toda $y' \in Y$ distinta de y. Más aún, $xBy \geq xBy'$ para toda $y' \in Y$ pues $x \in R_1(y)$. Luego y es también la única respuesta óptima a cualquier combinación convexa $x(\varepsilon) = (1-\varepsilon)x + \varepsilon x'$, para todo $\varepsilon \in (0,1]$, pues $R_1(y)$ es convexo y la función de pagos es lineal. Por estar hablando de un equilibrio que es perfecto en subjuegos, la respuesta óptima del seguidor a $x(\varepsilon)$ es también y. El pago para el líder cuando juega $x(\varepsilon)$ se calcula entonces como $(1-\varepsilon)xAy + \varepsilon x'Ay$ que está arbitrariamente cerca de $a(y)$ si tomamos ε suficientemente pequeño. Luego, si la respuesta de **J**2 en el camino al equilibrio (es decir, en busca de un equilibrio con la nueva estrategia anunciada por el líder) le deja a **J**1 un pago menor que $a(y)$, entonces podría cambiar su estrategia x por $x(\varepsilon)$ para un ε pequeño y obtener así un beneficio, contradiciendo la propiedad de equilibrio. Luego el pago del líder en el equilibrio Stackelberg es al menos $a(y)$. □

El Lema nos provee de procedimiento directo para calcular el equilibrio Stackelberg en el caso en que todas las estrategias de **J**2 son inducibles: Tomamos $y \in Y$ tal que $a(y)$ sea maximal, y supongamos que el líder anuncia que jugará $x \in R_1(y)$ de modo que se alcance $a(y)$, es decir, xAy es maximal. Entonces el seguidor responde de manera que **J**1 reciba en efecto dicho pago.

La prueba del lema explica la ventaja que tiene el líder, el jugador con la posibilidad de anunciar su estrategia. Tenemos que cualquier x tal que $a(y) = xAy$ es un punto extremo de $R_1(y)$, pues x maximiza una función lineal en el poliedro $R_1(y)$, y generalmente x pertenece a varias regiones $R_1(y')$. En ese caso, el seguidor es indiferente entre varias respuestas óptimas, pero nunca elige la que es favorable para el líder, dado que en otro caso éste induciría al seguidor a jugar lo que le conviene intercambiando x por $x(\varepsilon)$ y convirtiendo a y en la única respuesta racional.

Teorema 3.3. *Supongamos que toda estrategia $y \in Y$ es inducible. Entonces el pago del líder está determinado por*

$$L = \max_{y \in Y} a(y) = \max_{y \in Y} \max_{x \in R_1(y)} xAy \tag{3.2}$$

Más aún, ningun equilibrio de Nash conlleva a un mejor pago para el líder.

Demostración. El Jugador 1 no puede obtener un pago mayor que L pues el Jugador 2 siempre elige una respuesta óptima a la anunciada por él. Por otro lado, si el Jugador 1 obtendrá L jugando la estrategia x que realiza el máximo en (3.2), de acuerdo con el Lema 3.2.

Supongamos ahora que (p, q) es una estrategia que está en equilibrio de Nash para el juego simultáneo. Entonces cualquier estrategia pura que tenga asignada probabilidad positiva en q debe formar parte del conjunto de respuesta racional de

x, o dicho de otro modo, si $q_y > 0$ entonces $p \in R_1(y)$. Luego tenemos que para el pago Nash sucede:

$$\sum_{y \in Y}(pA)q_y = \sum_{y \in Y, q_y > 0}(pA)q_y \leq \sum_{y \in Y} a(y)q_y \leq \max_{y \in Y} a(y) = L$$

□

El problema será entonces cuando el seguidor tenga estrategias que no son inducibles. Esto indicará que existen otras estrategias que son equivalentes en resultado o que están débilmente dominadas por alguna estrategia diferente q, ya sea pura o mixta, de manera que $By \leq Bq$. Entonces siempre que y sea la respuesta óptima a algún $p \in \mathcal{P}$, también lo será q y todas aquellas estrategias puras indicadas por q con probabilidad positiva. Para comprender esta situación es que tenemos el siguiente lema:

Lema 3.4. *Una estrategia y es no inducible si y sólo si $By \leq Bq$ para algún $q \in \mathcal{Q}$ con $q_y = 0$.*

Demostración. Si el pago obtenido por el seguidor al jugar y es equivalente al de jugar q o está débilmente dominado, entonces q también será una respuesta óptima cuando y lo sea y por lo tanto no se cumple la unicidad en la definición de estrategia inducible.

Para probar la ida, supongamos que no existe $q \in \mathcal{Q}$ tal que $q_y = 0$ y $By \leq Bq$. Entonces podemos encontrar una estrategia $p \in \mathcal{P}$ tal que y sea la única respuesta óptima de $\mathbf{J}2$ mediante el teorema de dualidad para problemas de programación lineal, concluyendo que y es inducible. Consideremos el siguiente problema primal:

$$\begin{aligned} \max_{u,q} \quad & u \\ s.a. \quad & -\sum_{k \in Y \setminus \{y\}} Bkq_k + \mathbf{1}u \leq -By \\ & \sum_{k \in Y \setminus \{y\}} q_k = 1 \\ & q_k \geq 0, \forall k \in Y \setminus \{y\} \end{aligned}$$

El problema primal es factible, con cualquier $q \in \mathcal{Q}$ tal que $q_y = 0$ y u suficientemente negativo. El valor óptimo u^* será negativo, pues de otra manera tendríamos que $By \leq Bq$ contradiciendo nuestra hipótesis inicial.

El problema dual consta de hallar un vector no negativo $p = (p_1, ..., p_n)$ y un valor t de manera que $-pBy + t$ sea mínimo, sujeto a que $p_1 + ... + p_n = 1$. Dicho de otro modo, $p \in \mathcal{P}$ tal que $-pBk + t \geq 0$ para todo $k \neq y$.

Por teorema de dualidad [6], el problema dual asociado también será factible con valor óptimo igual al del primal, y este será negativo. Tendremos entonces que $xBy > t \geq pBk$ para todo $k \neq y$, es decir, y es la única respuesta racional para p, como queríamos. □

Si y está estrictamente dominada, entonces nunca será una respuesta racional, y tanto en el juego simultáneo como en el jerárquico el Jugador 2 nunca eligirá y pudiendo omitirla del juego.

El otro caso para que una estrategia y no sea inducible es cuando $By = By'$ para algún $y \neq y'$, dado que y e y' implican el mismo pago sea cual sea la estrategia que juegue el líder. Este caso es el que trataremos último.

La última opción para que y no sea inducible es que o bien sea equivalente en resultado a una estrategia mixta q (con probabilidad positiva asignada a por lo menos otras 2 estrategias puras) obteniendo $By = Bq$; o bien que y esté débilmente dominada por una estrategia q, con $By \lneqq Bq$. Para este último caso veremos a continuación que las regiones de respuesta óptima $R_1(y)$ tienen dimensión menor que la de X. Para esto primero recordemos un par de conceptos y resultados geométricos:

Una *combinación afín* de puntos $z^1, ...z^k$ en algún espacio Euclídeo es de la forma $\sum_{i=1}^{k} \lambda_i z^i$ para $\lambda_1, ..., \lambda_k \in \mathbb{R}$ tal que $\sum_{i=1}^{k} \lambda_i = 1$. La *cápsula afín* de los puntos $z^1, ...z^k$ es el conjunto de todas las combinaciones afines. Los puntos $z^1, ...z^k$ se dicen *afínmente independientes* si ninguno de ellos es una combinación afín de los otros, o equivalentemente, $\sum_{i=1}^{k} \lambda_i z^i = 0$ y $\sum_{i=1}^{k} \lambda_i = 0$ implican que todos los λ_i son nulos. Un conjunto convexo tiene dimensión k si y sólo si tiene a lo sumo $k + 1$ puntos afínmente independientes.

Como $\mathcal{P}$ es la cápsula convexa de n puntos afínmente independientes (las n estrategias puras del líder que podemos considerar como los vectores canónicos del espacio Euclideo) tiene dimensión $n-1$. Cualquier subconjunto P de $\mathcal{P}$ se dice que tiene *dimensión completa* si también tiene dimensión $n-1$. El siguiente resultado indica que si P tiene dimensión completa entonces es posible moverse dentro de P desde algún p una distancia pequeña en la dirección de cualquiera de los vectores unitarios p_i, los puntos extremos del simplex $\mathcal{P}$.

Lema 3.5. *Sea $p \in \mathcal{P}$ y $0 < \varepsilon \leq 1$. Entonces los vectores $(1 - \varepsilon)p + \varepsilon p_i$ son afínmente independientes para $1 \leq i \leq n$ y p es una combinación convexa de estos vectores.*

El siguiente resultado indica que los conjuntos de dimensión completa son aquellos que contienen un conjunto abierto en el sentido de la topología relativa a $\mathcal{P}$.

Notemos, para cada $\varepsilon > 0$, $B_\varepsilon(p)$ a la bola de radio ε alrededor de p considerando la norma euclídea.

Lema 3.6. *Sea P un subconjunto convexo de $\mathcal{P}$. Entonces son equivalentes:*

1. *P tiene dimensión completa;*
2. *Para algún $p \in P$ y algún $\varepsilon > 0$, los vectores $(1-\varepsilon)p + \varepsilon p_i$ pertencen a P para $1 \leq i \leq n$;*
3. *P contiene una bola abierta.*

Utilizando el tercer ítem de este lema podemos ver que $\mathcal{P}$ está cubierto por las regiones de respuesta óptima de dimensión completa. Observemos primero que

$$\mathcal{P} = \bigcup_{y \in Y} R_1(y) \tag{3.3}$$

pues cada $p \in \mathcal{P}$ tiene al menos una respuesta óptima $y \in Y$.

Proposición 3.7. *$\mathcal{P}$ es la unión de aquellos conjuntos de respuesta óptima $R_1(y)$ que tienen dimensión completa.*

Demostración. Sea $D = \{y \in Y \ / \ R_1(y) \text{ tiene dimensión completa}\}$. Consideremos el conjunto abierto $P = \mathcal{P} \setminus \bigcup_{y \in D} R_1(y)$ y supongamos que no es vacío, negando lo que queremos demostrar. Entonces para cualquier $y \in Y \setminus D$, el conjunto $D_y = D \setminus R_1(y)$ es también un abierto no vacío, pues si no tendríamos un entorno de P contenido en $R_1(y)$ contradiciendo el Lema 3.6. Repitiendo este argumento con $D_{y'}$ para $y' \in Y \setminus D$, obtenemos que $P \setminus \bigcup_{y \in Y \setminus D} R_1(y)$ es no vacío. Luego, $\mathcal{P} \setminus \bigcup_{y \in Y} R_1(y)$ es no vacío, contradiciendo lo observado en (3.3). □

El siguiente lema dice que la intersección de dos regiones de respuesta racional no puede tener dimensión completa a menos que induzcan pagos equivalentes para el seguidor.

Lema 3.8. *Si $R_1(y) \cap R_1(y')$ tiene dimensión completa entonces $By = By'$.*

Demostración. Consideremos, según el segundo ítem del Lema 3.6, un p y un $\varepsilon > 0$ de manera que $(1-\varepsilon)p + \varepsilon p_x$ pertenezca a $R_1(y) \cap R_1(y')$ para todo $x \in X$. Entonces, por el Lema 3.5, p es una combinación convexa de estos vectores y por lo tanto también pertenece a dicha intersección. Como y e y' son respuestas óptimas a p y a $(1-\varepsilon)p + \varepsilon p_x$ para todo x, tenemos que $pBy = pBy'$ y $((1-\varepsilon)p + \varepsilon p_x)\, By = ((1-\varepsilon)p + \varepsilon p_x)\, By'$. Esto implica que para todo $x \in X$, $p_x By = p_x By'$, es decir que los vectores columna By y By' coinciden componente a componente. □

El siguiente lema relaciona las estrategias inducibles con los conjuntos de respuesta óptima que tienen dimensión completa. El caso problemático sigue siendo el de los pagos equivalentes.

Lema 3.9. *Sea una estrategia pura del defensor $y \in Y$.*

1. *Si y es inducible entonces $R_1(y)$ tiene dimensión completa.*
2. *$R_1(y)$ tiene dimensión completa si y sólo si las únicas estrategias mixtas $q \in \mathcal{Q}$ que cumplen $By \leq Bq$ son aquellas para las que $p_{y'} > 0$ implica $By = By'$.*
3. *Si $By \neq By'$ para toda $y' \in Y \setminus \{y\}$ y $R_1(y)$ tiene dimensión completa, entonces y es inducible.*

Demostración. 1. Si y es inducible, el conjunto $\{p \in \mathcal{P} \ / \ pBy > pBy' \quad \forall y \neq y'\}$ es un abierto no vacío contenido en $R_1(y)$ que resulta de dimensión completa como consecuencia del Lema 3.6.

2. Supongamos que $R_1(y)$ tiene dimensión completa, y que existe una estrategia $q \in \mathcal{Q}$ tal que $By \leq Bq$ pero que $q_{y'} > 0$ y $By \neq By'$ para cierta $y' \in Y$. Con esto, $R_1(y) \subseteq R_1(y')$ pues siempre que y sea una respuesta óptima, y' también lo será. De esta manera $R_1(y) = R_1(y) \cap R_1(y')$ tiene dimensión completa y por el lema anterior $By = By'$, concluyendo un absurdo. Para la vuelta, supongamos que la desigualdad $By \leq Bq$ se cumple sólo cuando $q_{y'} > 0$ implica $By = By'$ y consideremos el juego sin ninguna de las estrategias puras del seguidor que cumplen $By = By'$. En este juego, no existe ninguna estrategia $q \in \mathcal{Q}$ que cumpla simultáneamente $By \leq Bq$ y $q_y = 0$, y por el Lema 3.4 resulta que y es inducible. Con esto, $R_1(y)$ tiene dimensión completa por el ítem 1.

3. Por el ítem 2, no existe $q \in \mathcal{Q}$ tal que $By \leq Bq$ y $q_y = 0$, y por lo tanto, nuevamente por el Lema 3.4, y resulta inducible. □

Concluimos con el siguiente teorema que generaliza lo establecido por el Teorema 3.3 pero excluye el caso en que dos estrategias distintas conllevan el mismo pago para el seguidor independientemente de la estrategia del líder.

Teorema 3.10. *Supongamos que $By \neq By'$ para todo par de estrategias puras distintas $y, y' \in Y$. Entonces el conjunto de todos los posibles pagos para el líder en un equilibrio Stackelberg es un intervalo $[L, H]$, donde*

$$L = \max_{y \in Y, y \text{ inducible}} \max_{x \in R_1(y)} xAy \quad H = \max_{y \in Y} \max_{x \in R_1(y)} xAy. \tag{3.4}$$

Demostración. Utilizando los Lemas 3.9. y 3.7. podemos escribir:

$$\mathcal{P} = \bigcup_{y \in Y,\ y \text{ inducible}} R_1(y) \tag{3.5}$$

Esto significa que el seguidor tiene una respuesta óptima inducible para cualquier estrategia del líder. Consideremos, como hicimos anteriormente, el "juego inducible" en el que restringimos el conjunto de estrategias del seguidor a las que son inducibles. Dada $p \in \mathcal{P}$ cualquier respuesta óptima pura y para p en el "juego inducible" es también una respuesta óptima en el juego original, pues $pBy' \leq pBy$ para toda y' inducible, y para cualquier respuesta óptima no inducible k de p en el juego original existe también una respuesta óptima que sí lo es por 3.5, y por lo tanto $pBk = pBy' \leq pBy$.

Luego, cualquier equilibrio Stackelberg del juego inducible será también un equilibrio del juego original. Por el Teorema 3.3 el pago del líder en el juego inducible es L como definimos en 3.4 y por lo tanto uno de los posibles pagos del líder en el juego original. Incluso si el Jugador 2 tuviera otras respuestas óptimas en el juego original, no podrían conllevar a un pago menor para el juego original por lo establecido en el Lema 3.2.

Por otro lado, el mayor pago al que puede aspirar el líder es claramente H como definimos en 3.4. Si pensamos en que el seguidor opta por la respuesta óptima que le conviene al líder y suponiendo que el líder juega p^* que realiza el máximo en la definición de H, con p^*Ay^*H, y el Jugador 2 juega y^* la respuesta óptima a p^*, entonces el pago del líder será H.

Que $L \leq H$ está claro. Nos resta ver que existe un equilibrio con pago S entre L y H cuando no vale la igualdad. Esto sucede cuando la respuesta óptima y^* de p^* no es inducible, y en ese caso, por el Lema 3.5, existe una respuesta óptima y para p^*, donde $p^*Ay \leq L$. Supongamos que el líder anuncia su estrategia p^* y que el seguidor juega una estrategia mixta que asigna probabilidad positiva sólo entre y e y^* de modo que el pago esperado para el líder sea S (que será una combinación convexa de p^*Ay y p^*Ay^*). Para cualquier anuncio de estrategia distinto de p^* por parte del líder, como el seguidor elige en su conjunto de respuesta racional, tendremos que p^* es de hecho el anuncio óptimo. Luego se trata de un equilibrio Stackelberg con pago S para el líder, como queríamos. □

Finalmente, observemos el caso en que $By = By'$, es decir, el de las estrategias que son equivalentes para el seguidor. Esto puede pensarse como una relación de equivalencia en el conjunto de estrategias puras del seguidor. Si dos estrategias son equivalentes entonces ninguna de las dos es inducible y por lo tanto el teorema anterior no puede aplicarse, porque en ese caso no vale el Lema 3.5. Sin embargo,

el Lema 3.7. sigue valiendo. De hecho, el pago del líder puede ser descrito considerando las estrategias cuyo espacio de respuesta óptima es de dimensión completa, sin la necesidad de pedir que sean inducibles, los cuales están bien caracterizados con lo establecido en el segundo ítem del Lema 3.9.

Llegamos así al último teorema que generaliza los dos anteriores, y establece los beneficios del líder en el juego jerárquico, independientemente de si alguna estrategia tiene otra equivalente o si define una región de respuesta racional de dimensión menor.

Teorema 3.11. *Sea* $D = \{y \in Y \ / \ R_1(y)$ *tiene dimensión completa*$\}$. *El conjunto al que pertenecen los pagos del líder es un intervalo* $[L, H]$ *donde*

$$L = \max_{y \in D} \max_{x \in R_1(y)} \min_{y' \in Y,\ By = By'} xAy' \quad H = \max_{y \in Y} \max_{x \in R_1(y)} xAy. \tag{3.6}$$

Además, el valor correspondiente a cualquier equilibrio de Nash se mantiene menor o igual que L.

Demostración. Para $y \in Y$, usemos la siguiente notación para la clase de equivalencia $[y] = \{y' \in Y \ : \ By = By'\}$. Según lo que discutimos, $[y]$ es un subconjunto de D siempre que $y \in D$.

Lo primero será ver que L definido como en (3.6) es el pago correspondiente al equilibrio Stackelberg para el líder que construiremos a continuación. Sean $y \in D$, $x \in R_1(y)$ e $y' \in [y]$ de manera tal que el máx − máx − mín de (3.6) se realice como $L = xAy'$. El líder anuncia su estrategia x y el seguidor responde con y', y a cualquier otra estrategia eligiendo la respuesta óptima que minimice la ganancia del líder. Entonces x es la mejor estrategia que puede anunciar el líder. Cualquier respuesta racional y' entre las estrategias de D es también una respuesta racional cuando se tienen en consideración todas las estrategias de Y, como indica la Proposición 3.7 y discutimos en el Teorema 3.10. De esta manera obtenemos un equilibrio Stackelberg con pago L para el líder.

Más aún, L es el menor pago posible si hablamos de un equilibrio, razonando de modo similar al del Lema 3.2. Como $R_1(y)$ tiene dimensión completa, contiene algún punto interior x', por el Lema 3.6. Además, cualquier combinación convexa $x(\varepsilon) = (1 - \varepsilon)x + \varepsilon x'$ para $\varepsilon \in (0, 1]$ es también un punto interior de $R_1(y)$, y la respuesta óptima del seguidor para $x(\varepsilon)$ es cualquier $y' \in [y]$. El mínimo de los pagos del líder sobre estos $y' \in [y]$ es $x(\varepsilon)Ay'$, que es una función continua para ε y está arbitrariamente cerca de L a medida que hacemos a ε tender a cero. Luego, el seguidor debe jugar de manera que el líder consiga un pago L en camino al equilibrio.

Como en el Teorema 3.10, es fácil ver que H es el pago más alto posible, y por lo tanto $[L, H]$ es el conjunto de todos los pagos posibles para estrategias en

equilibrio Stackelberg. Además, ningún pago correspondiente a un equilibrio de Nash puede exceder H.

Para probar que el juego simultáneo tiene equilibrios Nash con pagos de a lo sumo L para el Jugador 1, modificaremos el juego en dos pasos. Primero consideraremos el juego "de dimensión completa", de manera similar al "inducible" del teorema anterior, en el que Y es reemplazado por D, cuyos equilibrios siguen siéndolo en juego original. En un segundo paso, consideraremos el juego "factorizado" donde las estrategias puras del seguidor se corresponden con las clases de equivalencia $[y]$ para $y \in D$. Para cada $[y]$, el Jugador 2 tiene una única columna de pagos By, que es por definición la misma para cualquier estrategia en $[y]$ y por lo tanto la llamaremos $B[y]$.

Las columnas de pago para el líder en el juego "factorizado" se obtienen como combinaciones convexas de las columnas originales Ay' para $y' \in [y]$. En este sentido, consideraremos el "juego con matrices restringidas" donde el Jugador 1 elige x en $R_1(y)$ y el Jugador 2 juega mixto entre las estrategias puras $y' \in [y]$, con el pago de suma cero y columnas Ay' para el Jugador 1. Este juego tiene un valor, dado por

$$L_y = \max_{x \in R_1(y)} \min_{y' \in [y]} xAy'.$$

Claramente, L_y es el máximo real u tal que $xAy' \geq u$ para todo $y' \in [y]$ y $x \in R_1(y)$, donde lo último puede escribirse como $x \in \mathcal{P}$ y $x(By - Bz) \geq 0$ para todo $z \in Y \setminus [y]$. Tenemos el problema para el cual podemos plantear un programa lineal (LP):

$$\begin{array}{lll} \min\limits_{u,x} & -u & \\ s.a. & xAy' - u \geq 0 & y' \in [y] \\ & x(By - Bz) \geq 0 & z \in Y \setminus [y] \\ & x\mathbf{1} = 1, x \geq 0 & \end{array}$$

Donde $\mathbf{1}$ es una columna con n unos. El dual de este LP utiliza variables no negativas $s_{y'}$ para $y' \in [y]$ y w_z para $z \in Y \setminus [y]$ y una variable irrestricta t:

$$\begin{array}{ll} \max\limits_{t,s,w} & t \\ s.a. & \sum\limits_{y' \in [y]} Ay' s_{y'} + \sum\limits_{z \in Y \setminus [y]} (By - Bz) w_z + \mathbf{1}t \leq 0 \\ & -\sum\limits_{y' \in [y]} s_{y'} = -1 \end{array}$$

Consideremos una solución óptima del problema dual. Entonces t es igual al óptimo en el problema primal, es decir, $t = -u = -L_y$. Más aún, siempre que el Jugador 2 utilice la estrategia mixta en $\mathcal{Q}$ con probabilidades $s_{y'}$ para $y' \in [y]$, y cero en otro caso, entonces para cualquier $x \in R_1(y)$ el pago esperado para el Jugador 1 cumple:

$$\sum_{y' \in [y]} xAy's_{y'} \leq \sum_{y' \in [y]} xAy's_{y'} + \sum_{z \in Y \setminus [y]} (By - Bz)w_z \leq -x\mathbf{1}t = -t = L_y.$$

Esto significa que el Jugador 1 no puede obtener un pago mayor que L_y para ningún $x \in R_1(y)$ siempre que el Jugador 2 actúe según las probabilidades $s_{y'}$ para $y' \in [y]$; llamémoslas $s_{y'}^{[y]}$ pues dependen de $[y]$. En el juego "factorizado", la columna de pago $A[y]$ del Jugador 1 para $[y]$ está dada por

$$A[y] = \sum_{y' \in [y]} Ay's_{y'}^{[y]},$$

y por lo tanto

$$L_y = \max_{x \in R_1(y)} \min_{y' \in [y]} xAy' = \max_{x \in R_1(y)} xA[y]. \tag{3.7}$$

Por construcción, todas las columnas de pago $B[y]$ para el Jugador 2 en el juego "factorizado" son diferentes. Más aún, todas las regiones de respuesta óptima son de dimensión completa. Podemos entonces aplicar el Teorema 3.3 y por lo tanto cualquier equilibrio de Nash del juego "factorizado" el pago para el Jugador 1 es a lo sumo igual al pago del líder, que por (3.2), (3.7) y (3.6) es igual a

$$\max_{y \in D} \max_{x \in R_1(y)} xA[y] = \max_{y \in D} L_y = L.$$

Finalmente, cualquier equilibrio de Nash (x, q') del juego "factorizado" se traslada a un equilibrio de Nash (x, q) para el "juego de dimension completa", y por lo tanto al juego original, como sigue: El Jugador 1 juega x como antes, y el Jugador 2 elige $y' \in [y]$ para $y \in D$ con probabilidad $q_{y'} = q'_{[y]} s_{y'}^{[y]}$. Entonces el jugador 1 recibe el mismo pago esperado que antes, pues x es una respuesta óptima para q, y como $By' = B[y]$, cualquier $y' \in D$ tal que $q_{y'} > 0$ es una respuesta óptima para x, como queríamos. El pago Nash resultante para el Jugador 1 es a lo sumo L, como afirmamos. □

3.3. El equilibrio fuerte

Por último, y para finalizar este capítulo, conectaremos los conceptos de equilibrio que consideramos anteriormente al definir los juegos mediante los dos enfoques: el normal y el extensivo. La teoría clásica expuesta en el primer capítulo

está enmarcada dentro de la literatura de los juegos dinámicos y la teoría de optimización, como en [3]. Allí consideramos los pagos de los jugadores como costos a minimizar, dada su aplicación principal en problemas surgientes de la economía, aunque para ser consistentes con los últimos resultados y con el segundo capítulo, continuaremos con un enfoque de maximización de utilidades.

Consideremos un juego finito con dos jugadores, con conjuntos de estrategias mixtas $\mathcal{P}$ y $\mathcal{Q}$ para el líder y el seguidor respectivamente. Para $p \in \mathcal{P}$ y $q \in \mathcal{Q}$, denotemos con $J_1(p, q)$ al pago para el líder. En el primer capítulo definimos el pago Stackelberg para el líder como

$$S = \sup_{p \in \mathcal{P}} \min_{q \in R_2(p)} J_1(p, q). \tag{3.8}$$

Esta ecuación describe una mirada "pesimista" sobre todos los posibles equilibrios que podríamos obtener cuando el seguidor juega una respuesta óptima, dado que asumimos que éste hace lo que es peor para el líder. El conjunto $R_2(p)$ es compacto por ser un juego finito y por lo tanto podemos tomar mínimo en (3.8) en vez de un ínfimo. De todas maneras, $\min_{q \in R_2(p)} J_1(p, q)$ es una función discontinua de p, como hemos visto en el análisis del último ejemplo de la sección 1.4.

Esta discontinuidad significa para nosotros un problema, al igual que en toda la teoría de optimización, y se aborda mediante distintas aproximaciones de regularización que justifican tomar una solución que aproxima S arbitrariamente. (ver [17]).

El teorema que presentamos en esta sección establece que el pago Stackelberg S de 3.8 en el caso de un seguidor es exactamente el pago L definido en (3.6) aunque expresado de manera más concisa en (3.8). La principal diferencia es que en general, (3.8) *no* describe el comportamiento del seguidor en un equilibrio jerárquico secuencial como los estudiados en éste capítulo, donde el seguidor suele elegir la respuesta que es favorable para el líder en pos de obtener soluciones. Más aún, S es el menor pago posible para el líder para cualquier número de seguidores.

Teorema 3.12. *Consideremos un juego finito de dos jugadores con estrategias mixtas. Entonces el correspondiente juego jerárquico admite un equilibrio perfecto en subjuegos (p^*, q^*) con $q^* \in R_2(p^*)$ tal que $J_1(p^*, q^*) = S$. Cualquier otro pago para el líder es al menos S.*

Demostración. La principal dificultad es cuando el seguidor presenta indiferencia entre distintas estrategias, es decir cuando $R_2(p^*)$ no consta de un único elemento, implicando $J_1(p^*, q^*) \neq \min_{q \in R_2(p^*)} J_1(p^*, q)$.

Para cada p, el conjunto de respuesta óptima $R_2(p)$ es el conjunto de puntos fijos de una función continua adecuada $T_p : \mathcal{Q} \to \mathcal{Q}$, por ejemplo el funcional

definido por Nash [18]. Más aún, esta función es continua en p. Sea entonces F la función continua definida sobre $\mathcal{P} \times \mathcal{Q}$ como $F(p,q) = T_x(q) - q$ que toma valores en el espacio Euclídeo que contiene a $\mathcal{Q}$. Entonces $R_2(p) = \{q \ : \ F(p,q) = 0\}$ y la correspondencia

$$\bigcup_{p \in \mathcal{P}} (\{p\} \times R_2(p)) = F^{-1}(0)$$

define un conjunto cerrado (y por ende compacto) contenido en $\mathcal{P} \times \mathcal{Q}$.

Consideremos una sucesión $(p_n, q_n)_{n \in \mathbb{N}}$ de manera que

$$J_1(p_n, q_n) = \min_{q \in R_2(p_n)} J_1(p_n, q)$$

y que converja a S. Esta sucesión pertence a $F^{-1}(0)$ y tiene una subsucesión convergente con límite (p^*, q^*) en $F^{-1}(0)$. Como la función de utilidad del líder J_1 es continua, $J_1(p^*, q^*) = S$ como queríamos.

Hasta aquí hemos visto la existencia del equilibrio enunciado. Si (p, q) es otro equilibrio del juego jerárquico tal que $J_1(p, q) \leq S - \varepsilon$ para algún $\varepsilon > 0$, entonces para algún n tendremos que $J_1(p_n, q_n) > J_1(p, q)$ y el líder podría desviarse de p a p_n y obtener un pago más alto, contradiciendo la propiedad de equilibrio. Luego, S es el menor pago posible para el líder en una situación de equilibrio. □

Capítulo 4

Métodos de optimización para resolver juegos Stackelberg

4.1. Introducción y formulación inicial del problema

En este capítulo nos centraremos en el uso de las herramientas computacionales y los métodos que nos provee la investigación operativa para la resolución de juegos Stackelberg. Es sumamente importante encontrar respuestas eficientemente si queremos modelar problemas reales mediante la teoría de los juegos Stackelberg, y en muchas situaciones nos encontraremos con desafíos causados por el tamaño de los problemas o la cantidad de variables que debemos tener en consideración. Existen varios ejemplos en la literatura que dan fe que el uso de juegos Stackelberg para modelar problemas de seguridad se hace posible gracias a que éstos pueden verse como problemas de programación lineal entera mixta, que a pesar de ser teóricamente "difíciles" sus soluciones están a nuestro alcance mediante los modernos algoritmos y solvers con los que contamos.

La teoría de juegos nos provee de aproximaciones matemáticas para el despliegue de los recursos limitados que tiene una fuerza de seguridad en pos de maximizar su efectividad. Este tipo de investigaciones surgen en las últimas décadas, con la emergencia de la teoría de juegos computacional. En particular, el grupo de investigación cuyos trabajos son la base de lo que expondremos en este capítulo, ha concretado la aplicación de estas técnicas en una considerable variedad de situaciones reales.

La primer aplicación, *Assistant for Randomized Monitoring Over Routes* (ARMOR) ha demostrado ser una puesta en práctica exitosa de los algoritmos desa-

rrollados para juegos teóricos en el Aeropuerto Internacional de Los Angeles, en 2007 y en uso desde entonces. En particular, ARMOR utiliza estos modelos para aleatorizar los patrullajes de la policía y sus unidades caninas en puntos fijos del aeropuerto. La segunda, *Intelligent Randomization in Scheduling* (IRIS), fue desarrollada para el U.S. Federal Air Marshal Service, para distribuir los tenientes de a bordo en distintos vuelos estadounidenses, y está en uso desde 2009. La tercera, *Game-theoretic Unpredictable and Randomly Deployed Security* (GUARDS) para la U.S. Transportation Security Administration está siendo evaluada e impacta en más de 400 aeropuertos de Estados Unidos [26].

Veremos distintas formulaciones de los juegos Stackelberg como problemas de programación lineal entera mixta, y distintos algoritmos exactos y heurísticas para resolverlos. Nuestro interés final será modelar situaciones de seguridad, por eso es que nos referiremos a los juegos Stackelberg como *juegos de seguridad Stackelebrg*: el *defensor* será el líder, quien tome primero la decisión de cómo defender o proteger cierto lugar; y el seguidor se llamará *atacante*.

El defensor se encarga de proteger un conjunto de *targets* $T = 1, ..., |T|$ para lo cual cuenta con distintos recursos con los que diseñará recorridos o schedules, en los que protegerá uno o más targets. Aquí hay una distinción en los conjuntos de estrategias puras: el defensor distribuye la probabilidad sobre el conjunto de los posibles schedules, que llamaremos S, y el atacante elegirá atacar un target. Así, denotaremos $p = (p_s)_{s \in S}$ y $a = (a_t)_{t \in T}$ las estrategias mixtas del defensor y del atacante, respectivamente. Nos resta definir los pagos del juego, que dependerán de si el target que es atacado está siendo cubierto por el defensor o no. Supongamos que $U_1^c(t)$ es el pago para el defensor cuando atacan el target t y éste está cubierto y $U_1^d(t)$ el pago si está descubierto; análogamente, sean $U_2^c(t)$ y $U_2^d(t)$ los pagos para el atacante cuando elige atacar un target que está cubierto y descubierto respectivamente. Obviamente, $U_1^c(t) > U_1^d(t)$ y $U_2^c(t) < U_2^d(t)$, es decir, para el defensor es mejor que ataquen un target que está cubierto y para el atacante al revés.

Los pagos definidos sobre las acciones conjuntas serán entonces:

$$J_i(s,t) = \begin{cases} U_i^c(t) & \text{si } t \in s \\ U_i^d(t) & \text{si } t \notin s \end{cases}$$

Una vez hechas estas especificaciones y considerando a e como un vector de unos, podemos encontrar el vector de estrategia óptima para el defensor mediante el siguiente programa de dos etapas:

$$\begin{array}{ll} \text{máx}_{p,a} & \sum_{t \in T, s \in S} J_1(s,t) p_s a_t \\ s.a. & e^t p = 1,\ p \geq 0 \\ & a_t = argmax_{e^t a = 1, a \geq 0} \sum_{k \in T, s \in S} J_2(s,t) p_s a_k \end{array}$$

Una aclaración importante es que este modelo asume que en caso de empate, el seguidor elige la opción que le conviene al líder. Esto es, estamos trabajando con equilibrios de Stackelberg fuertes.

Ejemplifiquemos las situaciones que nos interesa modelar y resolver: supongamos que queremos planificar cómo defender un edificio con n entradas y que para ello contamos con un oficial que podemos apostar en una de estas. El atacante elegirá una de las entradas para ingresar al edificio. Tenemos que $S = T = \{1, ..., n\}$ y por lo tanto la cantidad de estrategias puras de la defensa son n. Si en cambio tenemos k oficiales para defender las entradas entonces las estrategias puras serían todos los subconjuntos de T de cardinal menor o igual que k y por lo tanto $|S| = \sum_{j=1}^{k} \binom{n}{j}$. Esto nos indica que el tamaño del problema, es decir el conjunto factible sobre el que tenemos que optimizar, se agranda exponencialmente a medida que agregamos targets o recursos.

En este punto es importante hacer la siguiente observación: para cada vector de estrategias del defensor p, la optimización que consideramos internamente para el atacante es un problema de programación lineal. Sabemos que el óptimo se alcanza en un vértice del simplex o región factible y esto se corresponde con alguna estrategia pura [6].

De esta manera, y sin pérdida de generalidad, podemos asumir que el atacante juega sólo estrategias puras y podemos transformar el problema de dos niveles que planteamos antes como un problema de programación entera mixta, gracias a que el adversario tiene una función de objetivo lineal. Para eso tenemos que considerar algunas variables adicionales y restricciones de holgura complementaria: una constante M suficientemente grande, una variable auxiliar d que representará la utilidad obtenida para el defensor y una segunda variable auxiliar q que será la ganancia óptima del atacante. El problema queda expresado de la siguiente manera:

$$
\begin{aligned}
\max_{p,a,d,q} \quad & d \\
s.a. \quad & \sum_{s \in S} J_1(s,t) p_s \geq d - M(1 - a_t) && t \in T \\
& \sum_{s \in S} J_2(s,t) p_s \leq q && t \in T \\
& \sum_{s \in S} J_2(s,t) p_s \geq q - M(1 - a_t) && t \in T \\
& e^t p = 1, \ p \geq 0 \\
& e^t a = 1, \ a \in \{0,1\}^{|T|}
\end{aligned}
$$

El primer conjunto de restricciones es una linearización de la ganancia del defensor dada la estrategia pura del atacante; el segundo indica que q es una cota superior de la ganancia del atacante para cualquiera de sus estrategias puras; y la tercera hace de q la utilidad óptima para el atacante para alguna estrategia pura.

4.2. Distintas formulaciones de JSS para programación lineal entera mixta

La forma general de encarar estos juegos es mediante el enfoque Bayesiano, que postula que alguno de los jugadores tiene información incompleta sobre los pagos del juego [20] [12]. En principio, a la hora de modelar cierta situación real, podemos pensar que la valoración del atacante sobre los distintos targets no es siempre la misma, y en ese caso la información del defensor sobre los pagos de su adversario es imperfecta. La forma en la que se agrega esto en el modelo es suponiendo que hay un conjunto L con los diferentes tipos de atacantes. Supondremos entonces que el atacante de tipo $l \in L$ es enfrentado con probabilidad π_l y en ese caso los pagos del juego vendrán dados por $J_i^l(s,t)$.

El problema de elegir la estrategia óptima para el líder en un juego Stackelberg es NP-hard para un juego Bayesiano con varios tipos de seguidores [7], es por eso que los investigadores continuan proveyéndonos de mejoras prácticas en los algoritmos. Por otro lado, el caso con un único seguidor admite un algoritmo polinomial (la idea desarrollada en el primer capítulo). El algoritmo DOBSS que presentaremos ahora es actualmente el más eficiente para éstos juegos [22] y está en uso actualmente para el planeamiento de los patrullajes de seguridad en el Aeropuerto Internacional de Los Angeles [26]. Este opera directamente con la representación compacta Bayesiana, asegurando mejoras en la velocidad de orden exponencial con respecto a métodos planteados inicialmente en [7], que requieren la conversión del juego Bayesiano a su forma normal mediante la transformación de Harsanyi. La característica notable del algoritmo es que, a diferencia de los algoritmos aproximados desarrollados hasta entonces, DOBSS nos returna una solución óptima exacta.

Diferenciando las posibles acciones para atacantes de tipo l con a^l podemos escribir la siguiente formulación (DOBSS) del problema mediante un modelo de

programación lineal entera mixta:

$$
\begin{array}{lll}
\max\limits_{p,a,d,q} & \sum\limits_{l\in L} \pi_l d^l & \\
s.a. & \sum\limits_{s\in S} J_1^l(s,t)p_s \geq d^l - M(1-a_t^l) & t\in T, l\in L \\
& \sum\limits_{s\in S} J_2^l(s,t)p_s \leq q^l & t\in T, l\in L \\
& \sum\limits_{s\in S} J_2^l(s,t)p_s \geq q^l - M(1-a_t^l) & t\in T, l\in L \\
& e^t p = 1, p\geq 0 & \\
& e^t a^l = 1, a^l \in \{0,1\}^{|T|} & l\in L
\end{array}
\qquad \text{(DOBSS)}
$$

En esta formulación la estrategia mixta del defensor es elegida de manera tal que maximice la utilidad esperada, ponderando sobre el conjunto de atacantes L. Esta representación es la base del algoritmo ERASER desarrollado en [27], que además considera un conjunto de restricciones sobre las estrategias factibles del defensor, algo que suele hacerse para adaptar el modelo a las realidades físicas de los espacios que quieren protegerse.

En otros trabajos similares podemos encontrarnos con dos formulaciones equivalentes. La correspondiente al problema (MJS) planteado en [27] es la siguiente:

$$
\begin{array}{lll}
\max\limits_{z,a,q} & \sum\limits_{l\in L}\sum\limits_{s\in S}\sum\limits_{t\in T} \pi_l J_1^l(s,t) z_{stl} & \\
s.a. & \sum\limits_{s\in S}\sum\limits_{t\in T} z_{stl} = 1 & l\in L \\
& \sum\limits_{s\in S} z_{stl} = a_t^l & t\in T, l\in L \\
& \sum\limits_{t\in T} z_{stl} = \sum\limits_{t\in T} z_{st1} & s\in S, l\in L \\
& 0 \leq q^l - \sum\limits_{s\in S}\sum\limits_{k\in T} J_2^l(s,t) z_{skl} \leq M(1-a_t^l) & t\in T, l\in L \\
& z_{stl} \geq 0 & s\in S, t\in T, l\in L \\
& e^t a^l = 1,\ a^l \in \{0,1\}^{|T|} & l\in L
\end{array}
\qquad \text{(MJS)}
$$

En este modelo se introdujeron las variables z_{stl} para representar mediante expresiones lineales el producto $z_{stl} = p_s a_t^l$ que aparece en la función objetivo bilineal del

defensor. Complementariamente, las tres primeras restricciones nos garantizan que es ése el comportamiento de z. Las restantes son equivalentes a las consideradas en la formulación anterior.

La tercer formulación no requiere el uso de la constante M ni de las restricciones de holgura complementaria correspondientes, y también está expresada en función de la variable $z_{stl} = p_s a_t^l$. Podemos encontrar esta formulación en MIPSG [14]

$$
\begin{aligned}
\max_{z,a} \quad & \sum_{l\in L}\sum_{s\in S}\sum_{t\in T} \pi_l J_1^l(s,t) z_{stl} & & \\
s.a. \quad & \sum_{s\in S}\sum_{t\in T} z_{stl} = 1 & l \in L & \\
& \sum_{s\in S} z_{stl} = a_t^l & t \in T, l \in L & \\
& \sum_{t\in T} z_{stl} = \sum_{t\in T} z_{st1} & s \in S, l \in L & \text{(MIPSG)} \\
& \sum_{s\in S} (J_2^l(s,t) - J_2^l(s,k)) z_{stl} \geq 0 & t,k \in T, l \in L & \\
& z_{stl} \geq 0 & s \in S, t \in T, l \in L & \\
& e^t a^l = 1,\ a^l \in \{0,1\}^{|T|} & l \in L &
\end{aligned}
$$

En este punto nos encontramos con que a pesar de que estas formulaciones resuelven el mismo problema no está claro cuál de las tres es más efectiva en la práctica. En [5] se presenta un trabajo teórico que muestra que la relajación lineal de la tercer formulación está contenida en la relajación lineal de la segunda, algo que sugiere que debería proveernos de un gap inicial menor y obtener soluciones en menor tiempo. En el mismo trabajo, además, se ve que la relajación de la tercer formulación tiene todos los puntos extremos enteros cuando hay un único tipo de adversario. Sin embargo, este último problema tiene más variables y más restricciones que la primer formulación, lo que hace que la resolución de su relajación lineal sea computacionalmente más compleja. Por otro lado, la primer formulación tiene menos variables y restricciones, pero el uso de las constantes M puede conducirnos a gaps iniciales muy grandes y a una convergencia lenta a la solución del problema entero.

4.2.1. Sobre la estructura de los pagos en aplicaciones de seguridad

Como vimos anteriormente, la cantidad de acciones posibles para el defensor, es decir, el cardinal del conjunto S sobre el que cual buscamos una estrategia mixta

óptima, se hace exponencialmente grande cuando disponemos de k recursos para posicionar en n objetivos. En realidad, como el atacante elegirá un target, nos interesa si la estrategia que juega el defensor cubre o no dicho target: el defensor tendrá una ganancia si el objetivo que es atacado está siendo patrullado y en caso contrario quién obtendrá una recompensa será el atacante. Si consideramos que los pagos del juego dependen de si hay un recurso asignado a un objetivo, y no de cómo estén asignados los demás recursos, entonces podemos asumir que las matrices de pago del juego son de la forma:

$$J_i^l(s,t) = \begin{cases} U_i^c(t) & \text{si } t \in s \\ U_i^d(t) & \text{si } t \notin s \end{cases}$$

donde $t \in s$ significa que en el schedule o recorrido s el defensor asigna un recurso al target t y viceversa. Asumiremos además que a lo sumo un recurso puede ser asignado simultáneamente a un objetivo.

4.3. Representación compacta

La ganancia esperada para el defensor se calcula como una esperanza sobre su vector de estrategia, obteniendo una ganancia cuando el schedule elegido cubre al target atacado y una pérdida cuando no lo cubre. Está claro que podemos escribir

$$\sum_{s \in S} J_1^l(s,t)p_s = \sum_{s \in S/t \in s} U_1^c(t)p_s + \sum_{s \in S/t \notin s} U_1^d(t)p_s$$

Definamos además la variable auxiliar $x_t = \sum_{s \in S/t \in s} p_s$ que represente la frecuencia total con la que cubrimos al target t si usamos la estrategia p_s. Podemos usar esto para expresar la función objetivo de la siguiente manera:

$$\sum_{s \in S} J_1^l(s,t)p_s = U_1^c(t) \sum_{s \in S/t \in s} p_s + U_1^d(t) \sum_{s \in S/t \notin s} p_s = U_1^c(t)x_t + U_1^d(t)(1 - x_t).$$

En efecto, usando la variable x_t de probabilidad total de cubrimiento del target t podemos representar el modelo (MJS) como sigue:

$$\begin{aligned}
\max_{p,a,d,q} \quad & \sum_{l\in L} \pi_l d^l \\
s.a. \quad & U_1^c(t)x_t + U_1^d(t)(1-x_t) \geq d^l - M(1-a_t^l) && t\in T, l\in L \\
& 0 \geq q^l - U_2^c(t)x_t - U_2^d(t)(1-x_t) \geq M(1-a_t^l) && t\in T, l\in L \\
& \sum_{t\in T} a_t^l = 1, \quad a^l \in \{0,1\}^{|T|} && l\in L \\
& \sum_{t\in T} x_t = m, 0 \leq x \leq 1 \\
& e^t p = 1, p \geq 0 \\
& x_t = \sum_{s\in S/t\in s} p_s && t\in T
\end{aligned}$$

Las restricciones en el cuarto renglón para el cubrimiento de x se corresponden con el hecho de que el defensor debe cubrir un total de m objetivos y se deduce de las correpondientes a p en la formulación original.

Si pudieramos deshacernos de las dos últimas restricciones entonces nos encontraríamos con un problema que no toma en cuenta el conjunto S que es nuestro mayor inconveniente computacional por su posible gran tamaño. Es decir, tendríamos sólo la variable x cuyo tamaño es el del conjunto T mientras que el tamaño de S es exponencialmente más grande, y entonces resultaría el siguiente problema que notaremos (MJS-JSS) que puede ser resuelto eficientemente:

$$\begin{aligned}
\max_{p,a,d,q} \quad & \sum_{l\in L} \pi_l d^l \\
s.a. \quad & U_1^c(t)x_t + U_1^d(t)(1-x_t) \geq d^l - M(1-a_t^l) && t\in T, l\in L \\
& 0 \geq q^l - U_2^c(t)x_t - U_2^d(t)(1-x_t) \geq M(1-a_t^l) && t\in T, l\in L \quad \text{(MJS-JSS)} \\
& \sum_{t\in T} a_t^l = 1, \quad a^l \in \{0,1\}^{|T|} && l\in L \\
& \sum_{t\in T} x_t = m, 0 \leq x \leq 1
\end{aligned}$$

La solución de éste problema será un vector $(x_t)_{t\in T}$ que representará las probabilidades óptimas con las que cubrir cada objetivo. La dificultad radicará en reconstruir, a partir de x, un vector $(p_s)_{s\in S}$ con las probabilidades óptimas con las que debemos asignar algún recurso a cada recorrido. Ésto será posible siempre y cuando no haya restricciones adicionales y si todas las posibles combinaciones de m objetivos son alguna acción factible del defensor.

Veamos cómo proceder de manera similar en las dos formulaciones equivalentes que presentamos anteriormente. Tanto en (DOBBS) como en (MIPSG), la función objetivo $\sum_{l\in L}\sum_{s\in S}\sum_{t\in T}\pi_l J_1^l(s,t)z_{stl}$ puede ser reescrita haciendo especial énfasis en la estructura de los pagos de la siguiente manera:

$$\sum_{s\in S}\sum_{t\in T} J_1^l(s,t)z_{stl} = \sum_{t\in T}\left(U_1^c(t)\sum_{s\in S/t\in s} z_{stl} + U_1^d(t)\sum_{s\in S/t\notin s} z_{stl}\right)$$

Para deshacernos del uso de S, consideremos las variables $w_{tl} = \sum_{s\in S/t\in s} z_{stl}$ sobre las estrategias s que cubren t y $\overline{w}_{tkl} = \sum_{s\in S/t\in s} z_{skl}$ la suma sobre las estrategias s que cubren el objetivo k fijo. Notemos primero que si tomamos $k = t$ tenemos $w_{tl} = \overline{w}_{ttl}$ y además $\sum_{s\in S/t\notin s} z_{skl} + \overline{w}_{tkl} = a_k^l$ para cualesquiera $t, k \in T$ y $l \in L$. Finalmente reescribamos:

$$\sum_{l\in L}\sum_{s\in S}\sum_{t\in T}\pi_l J_1^l(s,t)z_{stl} = \sum_{l\in L}\sum_{t\in T}\pi_l\left(U_1^c(t)w_{tl} + U_1^d(t)(a_t^l - w_{tl})\right)$$

$$= \sum_{l\in L}\sum_{t\in T}\pi_l\left(U_1^c(t) - U_1^d(t)\right)w_{tl} + \pi_l U_1^d(t)a_t^l$$

Así, podemos presentar una formulación equivalente del problema DOBBS utilizando las nuevas variables w_{tl}, a partir del cuales podremos reconstruir las originales z_{stl} siempre que no haya restricciones adicionales sobre los recorridos. Podemos encontrar una prueba de dicha equivalencia en [15]. Casi todas las restricciones se deducen inmediatamente de las planteadas en el problema original, tal vez la más complicada sea la correspondiente a la respuesta óptima del atacante: $0 \le q^l - \sum_{s\in S}\sum_{k\in T} J_2^l(s,t)z_{skl} \le M(1-a_t^l) \quad t \in T, l \in L$.

El procedimiento es el siguiente:

$$\sum_{s\in S}\sum_{k\in T} J_2^l(s,t)z_{skl} = \sum_{t\in T}\left(\sum_{s\in S/t\in s} U_2^c z_{skl} + \sum_{s\in S/t\notin s} U_2^d(t)z_{skl}\right)$$

$$= \sum_{t\in T}\left(U_2^c(t)\overline{w}_{tkl} + U_2^d(t)(a_k^l - \overline{w}_{tkl}\right)$$

$$= U_2^c(t)x_t + U_2^d(t)(1-x_t)$$

Con esto, la formulación del modelo DOBSS representado compactamente es:

$$\begin{aligned}
\max_{\overline{w},a,q} \quad & \sum_{l\in L}\sum_{t\in T}\pi_l\left(U_1^c(t)-U_1^d(t)\right)\overline{w}_{ttl}+\pi_l U_1^d(t)a_t^l & \\
s.a. \quad & \sum_{t\in T}a_t^l=1 & l\in L\\
& a_t^l\in\{0,1\} & t\in T, l\in L\\
& 0\le \overline{w}_{tkl}\le a_k^l & t,k\in T, l\in L\\
& x_t-1+a_k^l\le \overline{w}_{tkl}\le x_t+1-a_k^l & t,k\in T, l\in L\\
& \sum_{t\in T}x_t=m, 0\le x\le 1 & \\
& 0\le q^l-U_2^c(t)x_t-U_2^d(t)(1-x_t)\le M(1-a_t^l) & t\in T, l\in L
\end{aligned}$$

En el caso del MIPSG, la restricción de respuesta óptima es

$$\sum_{s\in S}(J_2^l(s,t)-J_2^l(s,k))z_{stl}\ge 0$$

y es aquí que el uso de las variables $\overline{w}_{tkl}$ cobra sentido:

$$\sum_{s\in S}(J_2^l(s,t)-J_2^l(s,k))z_{stl}=\left(U_2^c(t)w_{tl}+U_2^d(t)(a_t^l-w_{tl})\right)-\left(U_2^c(k)\overline{w}_{ktl}+U_2^d(k)(a_t^l-\overline{w}_{ktl})\right)$$

La formulación del problema MIPSG como un juego de Seguridad Stackelberg es:

$$\begin{aligned}
\max_{\overline{w},a} \quad & \sum_{l\in L}\sum_{t\in T}\pi_l\left(U_1^c(t)-U_1^d(t)\right)\overline{w}_{ttl}+\pi_l U_1^d(t)a_t^l & \\
s.a. \quad & \sum_{t\in T}a_t^l=1 & l\in L\\
& a_t^l\in\{0,1\} & t\in T, l\in L\\
& 0\le \overline{w}_{tkl}\le a_k^l & t,k\in T, l\in L\\
& x_t-1+a_k^l\le \overline{w}_{tkl}\le x_t+1-a_k^l & t,k\in T, l\in L\\
& \sum_{t\in T}x_t=m, 0\le x\le 1 & \\
& (U_2^c(t)-U_2^d(t))\overline{w}_{ttl}+U_2^d(t)a_t^l\ge (U_2^c(k)-U_2^d(k))\overline{w}_{ktl}+U_2^d(k)a_t^l & t,k\in T, l\in L
\end{aligned}$$

4.4. Modelos con incertidumbre observacional y de ejecución

Una vez que hemos desarrollado estos métodos surgen naturalmente nuevos desafíos: siempre que queremos modelar situaciones de la vida real nos encontra-

mos con que los agentes que intervienen en el proceso de toma de decisiones no son "perfectos" como requieren nuestras restricciones. La suposición general de "racionalidad" puede generar ciertas dudas pues no siempre nuestro oponente plantea la situación de la misma manera que nosotros. O lo hace, pero no sabe exactamente cuál es nuestra estrategia y comete errores al observarla, o suceden ciertos eventos inesperados que la modifican. Por ejemplo, un grupo de oficiales disponibles para un sector del aeropuerto podrían ser necesarios con urgencia en otro, teniendo que ser reasignados, o podrían aparecer nuevos recursos disponibles. De manera similar, el atacante podría no ver un oficial apostado en algun sector del aeropuerto, o confundir el traslado de un móvil con un patrullaje planificado. De este modo, tanto la ejecución de las estrategias del defensor como las observaciones del atacante para determinar estas estrategias están factibles de incertidumbre o "ruido".

Como analizamos en los juegos de información imperfecta, una estrategia que no tiene en cuenta esto podría no ser óptima en un ambiente incierto y por lo tanto es importante desarrollar técnicas que nos conduzcan a soluciones que tengan esto en consideración. Este tipo de investigaciones fueron desarrolladas para ser incluidas en las aplicaciones ARMOR e IRIS, en [23] y [31].

De esta manera, podemos pensar que en general el líder jugará una estrategia cosiderando tres tipos diferentes de incertidumbres, y que las dos primeras no disponen de una distribución de probabilidad a priori:

- incertidumbre sobre la respuesta óptima del seguidor debida a su racionalidad acotada, pudiendo éste tomar una estrategia que no maximice su utilidad;
- incertidumbre sobre la respuesta óptima del seguidor debida a los errores que podría cometer al observar la estrategia del líder;
- incertidumbre respecto de los pagos del seguidor modelada como diferentes matrices de pago con una suposición de distribución Bayesiana a priori (es decir, un juego Bayesiano).

Los algoritmos tratados hasta ahora modelan el tercer tipo de incertidumbre, en ellos habíamos considerado una distribucion a priori π sobre los diferentes tipos de atacantes, asignándole a cada uno una matriz de pago distinta. La optimalidad proporcionada por los modelos de la sección anterior falla si se presentan incertidumbres como las dos primeras, pudiendo ver una disminución impredecible en la ganancia del defensor.

En [23] los autores proponen nuevos algoritmos con los que obtener soluciones *robustas*, es decir, algoritmos que proporcionan soluciones efectivas a pesar de la incertidumbre sobre la elección del seguidor debido a su racionalidad acotada o

a la incertidumbre observacional. Presentaremos aquí los tres programas lineales enteros-mixtos (MILPs) desarrollados en dicho trabajo.

El primer MILP, BRASS (Bounded Rationality Assumption in Stackelberg Solver) tiene en consideración la racionalidad acotada del seguidor a la hora de optimizar la estrategia del líder. Supondremos que el seguidor no maximizará estrictamente su utilidad, sino que eligirá una estrategia ε-óptima, es decir, una estrategia cuyo pago difiere a lo sumo ε del pago óptimo. Dadas múltiples respuestas ε-óptimas, la aproximación robusta que utilizaremos constará de suponer que el seguidor elege la que conlleve el peor pago para el líder, no porque el seguidor vaya a hacer este razonamiento sino porque queremos obtener una solución robusta. Esto no es consistente con la suposición que hicimos en los otros algoritmos para resolver juegos Stackelberg, pero está basado en metodologías usuales de la optimización robusta, en la que los encargados de la toma de decisiones optimizan sobre el peor caso cuando se encuentran en una situación de incertidumbre [2][19], del mismo modo que afirma la teoría psicológica acerca de las decisiones que toman los humanos bajo incertidumbre observacional a la hora de elegir entre un conjunto discreto de acciones [25] [9].

El modelo que plantearemos a continuación esta basado en (MJS) presentado en la sección 4.2 y mantiene su notación. Agregamos las variables h_t^l, que representarán la estrategia óptima para un seguidor de tipo l, cuyo pago es q^l. Las variables a_t^l representan todas las estrategias ε-óptimas para un seguidor de tipo l. La segunda restricción permite ahora seleccionar más de una política para cada tipo de seguidor. La cuarta restricción hace de $a_t^l = 1$ para todas las acciones t tales que $q^l - \sum_{s \in S} J_2^l(s,t) < \varepsilon$, dado que en ese caso el término del medio es menor que ε y la desigualdad de la izquierda se satisface únicamente si $a_t^l = 1$.

Con esto, el algoritmo BRASS consistirá en resolver el siguiente MILP:

$$
\begin{aligned}
\max_{p,a,d,q,h} \quad & \sum_{l\in L} \pi_l d^l \\
s.a. \quad & \sum_{t\in T} h_t^l = 1 \\
& \sum_{t\in T} a_t^l \geq 1 \\
& \sum_{s\in S} J_2^l(s,t)p_s \geq q^l - M(1-h_t^l) && t\in T, l\in L \\
& \varepsilon(1-a_t^l) \leq q^l - \sum_{s\in S} J_2^l(s,t)p_s \leq \varepsilon + (1-a_t^l)M && t\in T, l\in L \quad \text{(BRASS)} \\
& \sum_{s\in S} J_1^l(s,t)p_s \geq d^l - M(1-a_t^l) && t\in T, l\in L \\
& h_t^l \leq a_t^l && t\in T, l\in L \\
& e^t p = 1, p \geq 0 \\
& a^l, h^l \in \{0,1\}^{|T|} && l\in L
\end{aligned}
$$

El segundo algoritmo, GUARD (Guarding Using Alpha Ranges in DOBSS), considera el caso en que el seguidor se encuentra con incertidumbre observacional a la hora de identificar la estrategia del líder. Para el desarrollo de este algoritmo se tuvieron en cuenta estudios psicológicos sobre la respuestas humanas bajo incertidumbre proveniente de observaciones limitadas o modificadas por el sujeto. A saber, cuando se les ofrece información limitada acerca de la frecuencia con la que un conjunto discreto de eventos sucede, los humanos tienden a pensar que dichos eventos ocurren bajo una distribución aleatoria uniforme, esto es también conocido como "la ignorancia precedente" (o *ignorance prior*)[9][25]. A medida que el observador adquiere más información actualizará sus creencias desde la ignorancia precedente hacia la verdadera distribución. Sin embargo, se ha mostrado que estas actualizaciones son generalmente subestimadas [25].

Dados estos estudios relacionados con las elecciones humanas, además de la política p_s del defensor, consideraremos la política que observa el seguidor inciertamente, y la notaremos p'_s. GUARD determinará la estrategia del seguidor en función de dos hipótesis, la nula y la alternativa. La nula será que el líder tome la acción s y la alternativa que no lo haga, y ambas serán construidas para cada acción $s \in S$. Basados en la hipótesis nula, la ecuación utilizada para construir la política dependerá de las cantidades k_n y k_a de elementos que pertencezcan a las particiones subjetivas que corresponden a la hipótesis nula y a la alternativa, respectivamente, de la siguiente manera: $p'_s = (k_n/k_a)^\alpha (p_s/(1-p_s))^{1-\alpha}$ (esto será 1

elemento en la nula y $n-1$ en la alternativa). Esta ecuación dependerá también de la estrategia del líder p_s y del valor $\alpha \in [0,1]$, que se decide en base al número de observaciones y determina qué tan cerca está el observador de la distribución uniforme en vez de la real.

Para determinar este α es importante hacer un análisis en base al problema real que se está intentando resolver. Los autores del trabajo desarrollaron una heurística basada en las elecciones del atacante contra DOBBS para el algorimo GUARD. En particular, se tomaron 40 decisiones hechas en base a 5 casos de observación del DOBSS, y se buscó la correlación entre estas elecciones y los casos no observados u observados de manera limitada en DOBSS. Se determinó dicho α considerando estos últimos casos representando la distribución uniforme y la observación no limitada como la real.

Finalmente, se sustituyó la variable p_s por p'_s en la restricción que asegura que el seguidor maximiza su pago, de manera acorde a lo que él cree. Luego, consideraremos nuevas respuestas basadas en p'_s y no en la original, y optimizaremos nuestra política contra la suya, actualizada según las observaciones establecidas. El nuevo MILP resulta:

$$
\begin{aligned}
\max_{p,a,d,q} \quad & \sum_{l \in L} \pi_l d^l & & \\
s.a. \quad & \sum_{s \in S} J_1^l(s,t) p_s \geq d^l - M(1-a_t^l) & t \in T, l \in L & \\
& 0 \leq q^l - \sum_{s \in S} J_2^l(s,t) p'_s \leq M(1-a_t^l) & t \in T, l \in L & \text{(GUARD)} \\
& e^t p = 1, p \geq 0 & & \\
& e^t a^l = 1, a^l \in \{0,1\}^{|T|} & l \in L &
\end{aligned}
$$

Observemos que, según la definición de p'_s, cuando $\alpha = 0$ el algoritmo GUARD es equivalente al DOBSS, pues se supone que el seguidor observa perfectamente la estrategia del líder.

El último algoritmo, COBRA (Combined Observability and Rationality Assumption), como su nombre indica, combina las mejoras de los dos algoritmos anteriores, y se obtiene simplemente incorporando las alteraciones hechas para BRASS y GUARD en un único MILP. Esto es, se incluyen los parámetros α y ε, y el resultado es muy similar al programa BRASS, con la única (pero sustancial) diferencia que en las restricciones 3 y 4 p_s es reemplazado por p'_s. La justificación para hacer esto es la misma que la presetada anteriormente. Además, como p'_s se obtiene a partir de p_s y del número de observaciones, este algoritmo es capaz de encontrar una política p_s que maximice la ganancia del líder considerando cómo el

adversario actualiza sus creencias. El MILP final se formula de la siguiente manera:

$$
\begin{aligned}
\max_{p,a,d,q,h} \quad & \sum_{l \in L} \pi_l d^l & & \\
s.a. \quad & \sum_{t \in T} h_t^l = 1 & l \in L & \\
& \sum_{t \in T} a_t^l \geq 1 & l \in L & \\
& \sum_{s \in S} J_2^l(s,t) p'_s \geq q^l - M(1 - h_t^l) & t \in T, l \in L & \\
& \varepsilon(1 - a_t^l) \leq q^l - \sum_{s \in S} J_2^l(s,t) p'_s \leq \varepsilon + (1 - a_t^l) M & t \in T, l \in L & \text{(COBRA)} \\
& \sum_{s \in S} J_1^l(s,t) p_s \geq d^l - M(1 - a_t^l) & t \in T, l \in L & \\
& h_t^l \leq a_t^l & t \in T, l \in L & \\
& e^t p = 1, p \geq 0 & & \\
& a^l, h^l \in \{0,1\}^{|T|} & l \in L &
\end{aligned}
$$

Para determinar el valor de α a utilizar en COBRA se utilizó la misma heurística que la desarrollada en GUARD. Esta vez, en cambio, se tomaron en cuenta las decisiones tomadas contra BRASS en vez de DOBSS. Como BRASS incorpora el ε, la estrategia usada es diferente que la obtenida por DOBSS y los jugadores interactúan de manera diferente. Como el ε es consistente para COBRA, se sigue que las actualizaciones estarán basadas en las políticas obtenidas por BRASS.

Podemos observar además, que el caso $\alpha = 0$ hace de COBRA equivalente a BRASS, si $\varepsilon = 0$ entonces es equivalente a GUARD, y si ambos son cero entonces es equivalente a DOBSS, la base de estos problemas de programación lineal.

Capítulo 5

La aplicación

5.1. Problemática

Hacia mediados de 2014 se acercaron al Instituto de Cálculo autoridades de la Policía de Seguridad Aeroportuaria que estaban enterados del trabajo de F. Ordoñez y su grupo de investigación en el que se aplicaron los modelos descritos en el capítulo anterior en el Aeropuerto de Los Angeles, en Estados Unidos.

A principios de 2005 el Presidente de la Nación, Néstor Kirchner (2003-2007), decidió, a través del Decreto 145/05, transferir orgánica y funcionalmente a la Policía Aeronáutica Nacional (PAN) desde el ámbito del Ministerio de Defensa a la esfera del Ministerio del Interior, constituyendo la Policía de Seguridad Aeroportuaria (PSA) e incorporándola al sistema de seguridad interior establecido en la Ley 24.059 de Seguridad Interior. Asimismo, mediante aquel decreto, ordenó la Intervención de la PSA a través de la designación de un funcionario civil a los efectos de que ejerza las competencias, facultades y atribuciones propias del Director Nacional de la ex PAN y efectúe y/o proponga las reestructuraciones que considere pertinentes a los fines de proceder a la normalización del funcionamiento de la citada institución. Por su parte, mediante el Decreto 147/05, el Poder Ejecutivo Nacional designó a Marcelo Fabian Sain como Interventor de la PSA. [24]

El 31 de mayo de 2006, el Congreso Nacional sancionó la Ley 26.102 de Seguridad Aeroportuaria a través de la cual se convalidó la creación de la PSA. A los pocos días, esa ley fue promulgada por el Poder Ejecutivo Nacional. En verdad, el origen de la PSA constituyó una expresión más de las tendencias estructurales que signaron el gobierno y la gestión de los asuntos de la seguridad pública desde el retorno a la democracia a fines de 1983.

La decisión gubernamental de conformar esta nueva policía especializada dio

lugar a un proceso institucional que resultó novedoso por tres razones significativas. En primer lugar, porque a esta novel policía se la estructuró sobre la base de nuevos parámetros doctrinales, organizacionales y funcionales que fueron elaborados y formulados por el equipo designado para estar a cargo de este proceso, que llamaron "la Intervención". En segundo término, porque, de acuerdo con ese modelo policial, constituyó la primera experiencia institucional de una policía con mando civil, es decir, con una estructura de dirección superior y administración general conducida por funcionarios especializados en asuntos de seguridad pública pero sin estado policial. Y, en tercer lugar, porque significó la creación en el ámbito federal de la primera institución policial en democracia.

Sobre la base de estos parámetros, la Intervención delineó el perfíl orgánico-funcional de la PSA, es decir, el tipo de institución policial que se pretendía construir y que constituía un modelo novedoso. Dicho perfil tenía particularmente en cuenta lo que el equipo de conducción de la Intervención denominaba el "modelo tradicional de policía". Para Sain así como para su principal colaborador, Germán Montenegro, el modelo de policía que debía instituirse en la PSA debía contrastar sutancialmente con el conjunto de anacronismo y defectos doctrinarios, organizacionales y funcionales propios de las policías tradicionales que convertían a estas instituciones en organizaciones colapsadas. La interpretación que ellos tenían acerca de las policías tradicionales era clara: creían que eran instituciones obsoletas para abordar el conjunto de desafíos que se imponían frente a la complejidad y diversidad de las problemáticas criminales que debían afrontar. En concreto, para ellos, la PSA debía superar esos vicios y deficiencias institucionales y, en torno de este contraste, diseñaron el modelo de policía que los interventores formularon como proyecto institucional. En esa perspectiva, la nueva policía debía estar exclusivamente especializada en el desarrollo de las labores de prevención e investigación de los delitos en el ámbito aeroportuario, descartando de antemano el desenvolvimiento de cualquier otro tipo de tareas ajenas al control criminal. Dicho de otro modo, la PSA debía ser una organización especializada en el control del delito [24].

Sobre estos lineamientos es que el Director General de Planificación nos propuso desarrollar un software que mediante automatización y aleatorización abordara una redistribución de los distintos recursos de patrullaje de la PSA en los sectores públicos del Aeropuerto Internacional Ministro Pistarini, sito en Ezeiza. Para eso mantuvimos un conjunto de reuniones iniciales con integrantes del Departamento de Inteligencia Criminal Aeroportuaria (DICA), con quienes identificamos la problemática incial: los datos y recursos disponibles, el reconocimiento de la infraestructura que debíamos proteger, el funcionamiento sectorizado del aeropuerto en términos de servicios y demandas habituales, y distintas especificidades técnicas concernientes a la aplicación.

Se trata básicamente de dos terminales (A y C) de arribos y partidas donde apostar oficiales fijos, y dos estacionamientos y un sector de la autopista Richieri que constituyen la jurisdicción de la PSA donde asignar patrullajes móviles. El número de oficiales y patrullas varía a cada día y turno (diurno y nocturno) y por lo tanto debíamos dotar al modelo de esta flexibilidad, entre otras. Aún así, este número es siempre inferior al necesario para proteger todo el aeropuerto y es por eso que optimizar las locaciones significa aprovechar de la mejor manera los recursos escasos.

La PSA contaba con un registro georreferenciado de los delitos sucedidos en cada uno de estos sectores, con datos adicionales como el horario o si se trató de un robo o un hurto. Dicha información estaba en desuso pero era sumamente útil, pues nos daba la posibilidad de diferenciar las distintas zonas en términos de qué tan frecuentes eran los delitos y de qué tipo eran. En su gran mayoría, hablabamos de delitos menores: carteristas y hurtos provenientes de descuidos, robos en sectores alejados, y sujetos que se apostaban habitualmente a ofrecer servicios de taxi de manera ilegal. De todos modos, se registraron también robos de vehículos estacionados dentro del aeropuerto y cerca de éste en el sector de la autopista antes mencionado.

El hecho de que los recursos disponibles no sean suficientes para cubrir todos los sectores deseables del Aeropuerto, y que éstos además tengan distintas valoraciones y vulnerabilidades, sumado a que los procedimientos de la PSA en materia de prevención del delito se llevan a cabo todo el tiempo y los posibles atacantes pueden observar cómo lo hace (interacción líder-seguidor), hacen que un Modelo de Seguridad Stackelberg sea ideal para atacar esta problemática.

5.2. El modelo

Conseguir modelar fielmente una situación real mediante teoría de juegos es una tarea complicada, pues es necesario reflejar en ecuaciones y valores cuantitativos muchos aspectos que hasta entonces no habían sido planteados de esa manera. Para esto, la interacción con quienes conocen el funcionamiento habitual del Aeropuerto y se encargan de distribuir los recursos apelando a su juicio y experiencia es crucial.

Es por eso que en conjunto con personal de la Unidad Opearacional de Seguridad Preventiva (UOSP) y del Departamento de Inteligencia Criminal Aeroportuaria se definieron los objetivos o targets en los que distribuir el despliegue policial: el Aeropuerto estaría dividido en tres *zonas*, la terminal A, la C y el estacionamiento; y cada una de ellas estaría organizada en distintas *cuadrículas*. Además, para ser consistentes con la jornada laboral de los trabajadores, planificaríamos dos *turnos*

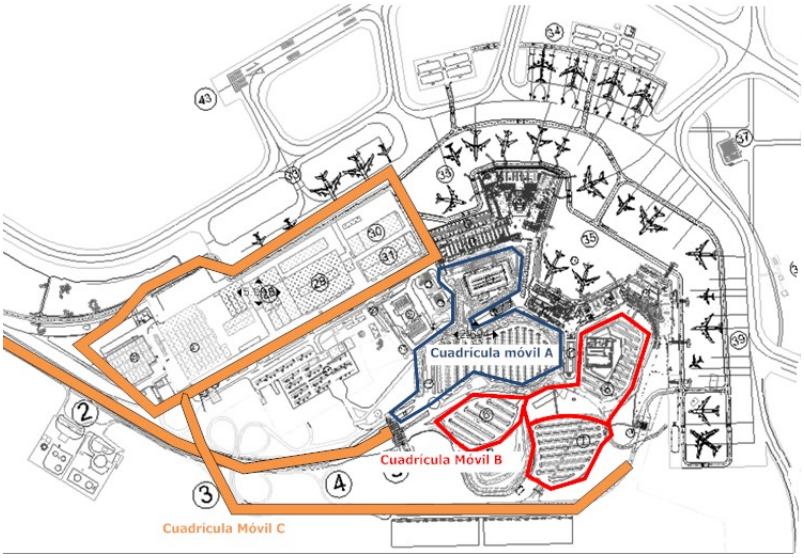


Figura 5.1: Plano General de Ezeiza, zona de patrullajes móviles

(diurno y nocturno) a los que asignar el personal y los recursos disponibles en su comienzo. Este número es muy variable pues existen varios factores que afectan: autoridades políticas que requieran mayor seguridad, allanamientos u operativos que impliquen un despliegue numeroso de agentes o tareas de patrullaje en otros sectores privados, que no forman parte de los que planificamos por ahora. De todos modos, a fines de entender las dimensiones del problema, podemos pensar que habitualmente se cuenta con un número de entre 10 y 20 efectivos para todas estas zonas.

A su vez, cada turno estaría subdividido en *franjas horarias* cuya duración es modificable por el usuario, aunque por default sería considerada como un período de una hora reloj. De esta manera, el conjunto de targets sobre los que optimizar el despliegue de los recursos constará de estas cuadrículas y el modelo se resolverá al principio de cada turno, para cada área, cuando se hayan cargado los recursos disponibles.

En las las figuras 5.1 y 5.2 podemos observar cómo quedaron definidas las cuadrículas del estacionamiento y de la terminal A.

El siguiente desafío es la determinación de las matrices de pago: no es equivalente para el atacante, en términos monetarios, robar una cartera o una valija, que una camioneta del estacionamiento; tampoco es lo mismo para la PSA, en términos de impacto social, un hurto en el hall principal que en un sector alejado del aeropuerto. De todas maneras estas dos situaciones son levemente distintas, pues la primera significa dos atacantes de distinto tipo, mientras que la segunda se refiere

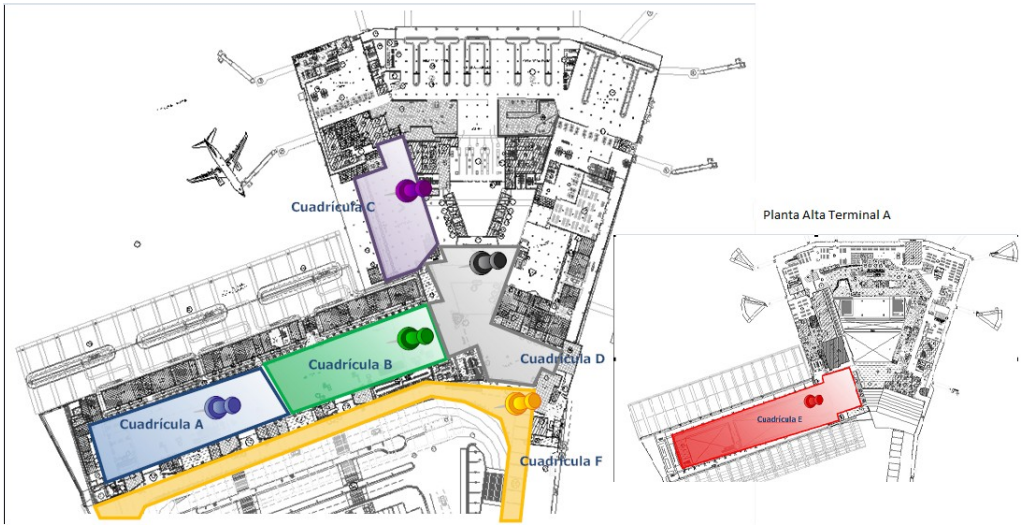


Figura 5.2: Cuadrículas para la Terminal A

a dos valoraciones distintas para diferentes cuadrículas. Ésto nos sugiere naturalmente el uso de modelos de tipo Bayesianos y la consideración de las frecuencias con las que se acontecen los delitos (datos que están a nuestra disposición). También tuvimos en cuenta el flujo de pasajeros habitual en los distintas cuadrículas (pues en principio, a mayor densidad de personas es mayor la probabilidad de que se cometa un hurto), y lo incorporamos como una variable de la que dependen los pagos del juego junto con los datos de ocurrencia de delitos, ponderados respecto a qué tan recientes fueron. De esta manera quedan determinadas las matrices de pago para cada zona y para cada franja horaria, que dependen linealmente del flujo de pasajeros y cuadráticamente de la cantidad de delitos cometidos en cada cuadrícula, teniendo en cuenta las ponderaciones definidas por la DICA: máxima ponderación para una semana hacia atrás, media para una semana a un mes, y menor ponderación para los delitos ocurridos un mes hacia atrás.

Es importante aclarar que la actualización de los datos de delitos se hace automáticamente mediante la integración de nuestro software con el sistema de la PSA que contiene dicha información, que está diferenciada por horario de ocurrencia del delito y ubicación. Como las matrices de pago son en principio distintas para cada horario, en cada caso nuestro sistema plantea un modelo de programación lineal diferente.

El modelo que utilizamos es similar al MJS que definimos en la sección 4.2. Presupusimos dos tipos de atacantes diferentes, pero éstos se esperan en distintas zonas del aeropuerto para las que resolvemos modelos independientes, con lo cual asumimos un solo tipo de seguidor.

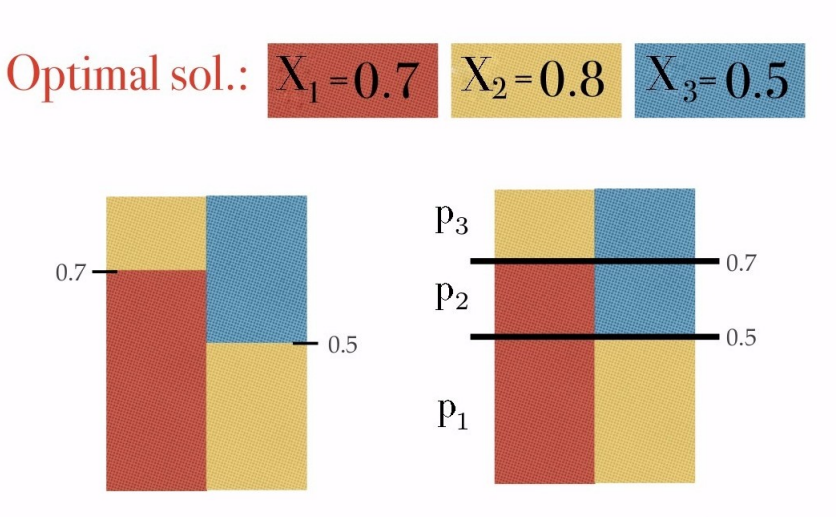
Las variables del modelo son x_t para $t \in T$ que nos indicarán las probabilidades óptimas con las que cubrir cada target t, a_t un vector de ceros con un uno indicando el objetivo que atacará el seguidor y, por último, q y D son variables auxiliares que indicarán, según la primer y segunda restricción, los pagos óptimos del atacante y defensor, respectivamente.

Los parámetros del modelo son los pagos del juego. En las matrices $U_1^c(t)$ y $U_1^d(t)$ están los pagos del defensor cuando el target t está cubierto y descubierto, respectivamente; del mismo modo, U_2 representa los pagos del atacante.

$$
\begin{aligned}
\max_{x,a,D,q} \quad & D \\
s.a. \quad & U_1^c(t)x_t + U_1^d(t)(1-x_t) \geq D - M(1-a_t) && t \in T \\
& 0 \geq q - U_2^c(t)x_t - U_2^d(t)(1-x_t) \geq M(1-a_t) && t \in T \\
& \sum_{t \in T} a_t = 1, \quad a \in \{0,1\}^{|T|} \\
& \sum_{t \in T} x_t = m, 0 \leq x \leq 1 \\
& D, q, M \in \mathbb{R}
\end{aligned}
$$

Una vez que resolvimos el modelo debemos reconstruir, a partir de x, un vector $(p_s)_{s \in S}$ con las probabilidades óptimas con las que debemos asignar algún recurso a cada schedule. Para esto, observemos primero que, según la cuarta restricción, $\sum_{t \in T} x_t = m$ y veamos con un ejemplo cómo hacer esta reconstrucción.

Supongamos que estamos en el caso particular en que tenemos 2 recursos y 3 objetivos que cubrir. Supongamos que corrimos el modelo y obtuvimos como solución óptima las siguientes probabilidades de cubrimiento: $x_1 = 0,7$, $x_2 = 0,8$ y $x_3 = 0,5$. Procedemos de la siguiente manera, ubicando en columnas de alto 1 los valores obtenidos, "cortando" las columnas para no excedernos:



Según la cuarta resticción, debemos sumar exactamente 2, por lo tanto completamos 2 columnas. Ahora seccionamos como en la figura, obteniendo 3 regiones: la correspondientes a cada uno de los schedules posibles: $p_1 = 0,5$ será la probabilidad de cubrir $s = \{1, 2\}$, $p_2 = 0,2$ la correspondiente a $s = \{1, 3\}$ y $p_3 = 0,3$ la de cubrir $\{2, 3\}$. De esta manera reconstruimos un vector óptimo p que satisface las probabilidades óptimas con las que cubrir cada recorrido.

El modelo es resuelto, y mediante el procedimiento descrito reconstruimos p: ahora tenemos la información suficiente para definir computacionalmente una distribución de probabilidades óptima con la cual designar un tiempo de cubrimiento concreto a cada recorrido para el turno que se esté programando. Con estos tiempos concretos, generaremos una distribución de los recursos a cada cuadrícula para todas las franjas horarias, para luego escribir un reporte de turno que será leído y ejecutado por el encargado de la designación.

5.3. El software

El software debía ser capaz de recibir los parámetros sobre los cuales definimos el problema y a su vez devolver al usuario un reporte final que indicara cómo distribuir el personal en los sectores del aeropuerto. El modelo Stackelberg explicitado en la sección anterior forma parte del *back end* del software que fue desarrollado como objetivo final de este proyecto. Los componentes tecnológicos utilizados en la construcción de este software fueron C# como lenguaje de programación, sobre un entorno Microsoft .NET Framework 4.2 y tecnología Winforms para la construcción de aplicaciones de cliente pesado.

El módulo *front end* constituye la interfaz gráfica a través de la cual el usuario

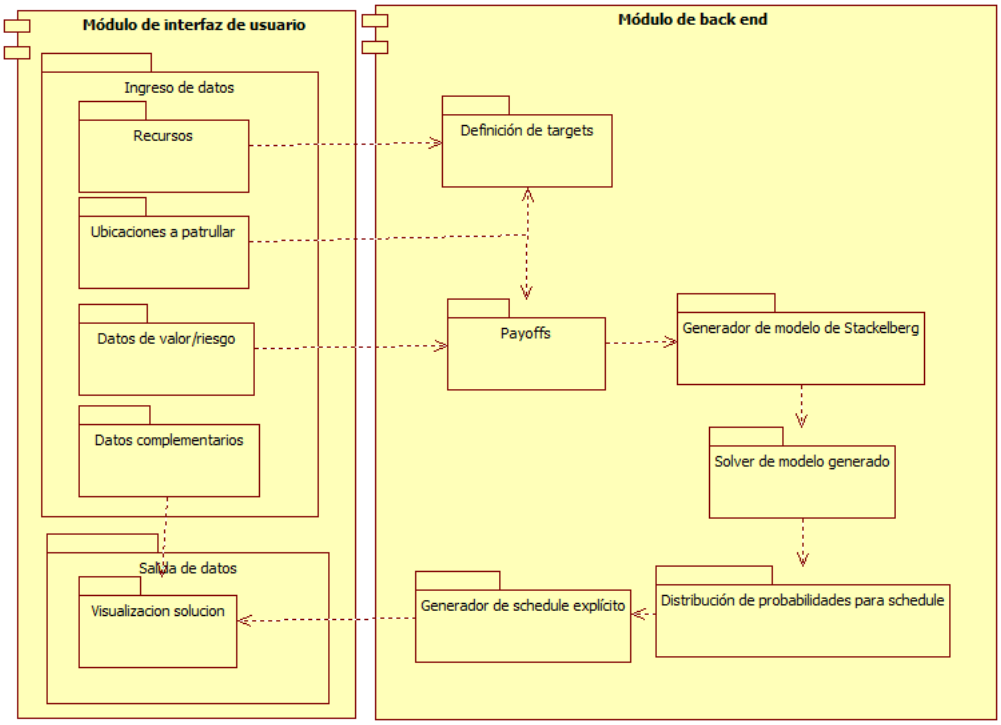


Figura 5.3: La Arquitectura del Software

de la aplicación interactúa con el sistema, en el que se distinguen naturalmente el ingreso y la salida de datos. Para dicha interfaz, se utilizó tecnología ASP.NET (Webforms).

El personal encargado de la planificación asignará la cantidad de oficiales disponibles al comienzo del turno a cada área del aeropuerto y modificará, si es necesario, el flujo de pasajeros en las zonas del aeropuerto para las diferentes franjas horarias, que puede ser bajo, medio o alto. Ambas funciones están disponibles en la aplicación. Complementariamente, el sistema cuenta con una base de datos con la información georreferenciada de los delitos ocurridos. Para ello, fue definido el siguiente modelo de datos, utilizando Entity Framework Versión 6, bajo el esquema Code First. Esto permite contar con un modelo de persistencia eficiente, que facilita las tareas de programación. El modelo fue utilizado "detrás" de todas las pantallas como mecanismo de almacenamiento de datos. La base de datos que recibe la información es SQL Server Compact Edition Versión 4 y está integrada al sistema SIGIPOL en el que se llevaba un registro digital de los delitos, lo que permite una actualización inmediata de nuestro sistema desde la base de datos de la PSA.

El componente de salida de datos es responsable de la visualización de la solución generada por el sistema, el reporte de turno que explicita a qué cuadrículas del aeropuerto asignar los recursos disponibles.

El módulo *back end* se encarga de traducir la información ingresada por el usuario a un modelo de Juego de Stackelberg. Este modelo es resuelto por un Solver de programación lineal entera, es decir, componentes de software que contienen algoritmos capaces de resolver modelos de programación lineal entera, encapsulados como librerías. En nuestro caso, utilizamos SCIP en su versión 3.1.0 [1]. La solución generada por el solver es una distribución de probabilidades. A partir de esta solución, se genera un modelo explícito mediante un sampleo de dicha distribución, con asignación de recursos concretos a ubicaciones de patrullaje concretas. Este resultado es informado al módulo Salida de datos, de forma tal de que éste construya una visualización de la misma para el usuario.

Se ajustó el modelo para su ejecución por franjas horarias. Esto es, la aplicación realiza actualmente una ejecución para cada franja horaria, tomando en consideración los recursos disponibles para esta franja. Se procede de esta manera en forma análoga a lo realizado en el Aeropuerto de Los Angeles en los trabajos tomados como referencia.

El software debe recibir por única vez la carga de los datos de zonas y cuadrículas del Aeropuerto, en conjunto con imagenes y mapas para una visualización amigable tanto de la planificación previa como del reporte de turno generado. Esta característica es notable, puesto que la misma aplicación, con una carga inicial

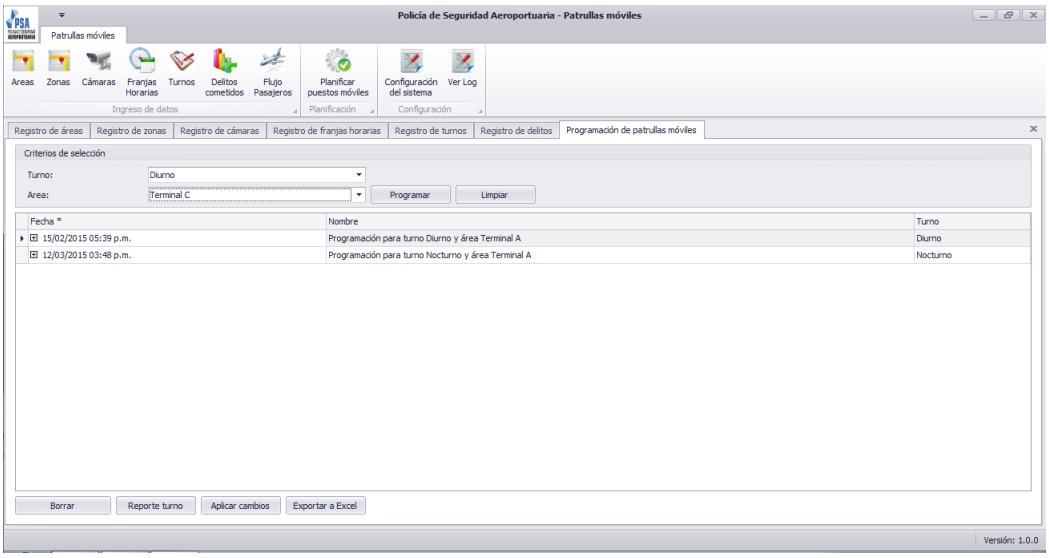


Figura 5.4: La pantalla principal de la aplicación

pertinente, podría servir para realizar un trabajo similar en sectores del Aeropuerto que no fueron considerados en esta primera instancia o, mejor aún, en otros Aeropuertos.

Conclusiones e investigaciones futuras

A lo largo de este trabajo hemos tratado la Teoría de Juegos clásica, analizando ambas formas de representar el juego: normal y extensiva. En ambas hemos hablado tanto de los juegos simultáneos como jerárquicos, y analizado particularmente los juegos Stackelberg y los equilibrios que hemos considerado como soluciones óptimas para el líder tanto en estrategias puras como mixtas. Discutimos sobre las ventajas de cada forma y sobre la existencia de soluciones, además de los métodos para encontrarlas. Diferenciamos los casos de información completa e incompleta, y el caso particular de los juegos Bayesianos, alcanzando suficiente generalidad teórica sobre la que sustentarnos. En el tercer capítulo relacionamos los primeros capítulos y convergimos a la solución sobre la que desarrollaríamos la teoría de optimización: el equilibrio Stackelberg fuerte.

El cuarto capítulo trata los temas de investigación en curso, muchos de los trabajos tomados como referencia son actuales, combinando un análisis teórico de los métodos planteados para encontrar las soluciones, distintas formulaciones de los problemas y un impacto social considerable debido a la eventual aplicación de la teoría en el mundo real. En esta línea hay todavía una gran variedad de problemas abiertos, cuando imponemos más restricciones en los problemas la existencia de soluciones ya no está garantizada y mucho menos la eficiencia de los algoritmos con los que las buscamos.

La teoría desarrollada en cuanto a la aplicación de modelos Stackelberg para resolver problemas de seguridad excede claramente lo que fue necesario para diseñar el motor de la aplicación llevada a cabo para el Aeropuerto de Ezeiza, posibilitando mejoras sustanciales en su funcionamiento siempre que se pueda hacer un relevamiento satisfactorio sobre su rendimiento. El primer paso será extender su uso a los sectores privados del aeropuerto, y la agregación de tareas de fondeo y cámaras. En una segunda instancia, y supeditados a un estudio pertinente, el agregado de incertezas en el modelo vía una relajación de la racionalidad de los agentes y distorsión de la observación de los seguidores de la estrategia del defensor.

Actualmente, el software se encuentra instalado y en etapa de recopilación de datos, tanto de los delitos cometidos y del índice de imputabilidad como de los reportes entregados por la aplicación. En cuanto estén a nuestra disposición los datos de los delitos podremos analizar en una primer instancia si éstos se vieron modificados de alguna manera. Luego, hay que cruzar estos datos con los reportes generados los días que acotencieron los delitos, para poder analizar si había o no asignado un recurso a dicha cuadrícula, estableciendo las frecuencias en ambos casos. En conjunto con los datos de imputación podremos determinar la existencia y orden de la incerteza de ejecución que afecta a nuestro modelo, para incorporarla a este y asemejarlo más fielmente a la realidad.

Más allá del impacto cuantitativo en la reducción de los delitos, el software significó un avance administrativo muy importante para la PSA y la adaptabilidad con la que cuenta el sistema para ser utilizado en otros aeropuertos resulta prometedora para trabajos futuros.

Bibliografía

[1] T. Achterberg. Scip: Solving constraint integer programs. *Mathematical Programming Computation*, 1(1):1–41, 2009.

[2] M. Aghassi and D. Bertsimas. *Robust game theory.* Math Program, 2006.

[3] T. Basar and G.J. Olsder. *Dynamic Noncooperative Game Theory.* Classics in Applied Mathematics, 1995.

[4] M. Breton, A. Alj, and A. Haurie. Sequential stackelberg equilibria in two-person games. *Journal of Optimization Theory and Applications*, 59:71–97, 1988.

[5] C. Casorrán-Amilburu, B. Fortz, M. Labbé, and F. Ordóñez. Novel formulations for stackelberg security games. *Technical report, Départment d'Informatique*, 2015.

[6] V. Chvátal. *Linear Programming.* W.H. Freeman and Co., 1983.

[7] V. Conitzer and T. Sandholm. Computing the optimal strategy to commit to. *EC*, 2006.

[8] T.S. Ferguson. *Game Theory.* UCLA, 2001.

[9] R. Fox and Y. Rottenstreich. Partition priming in judgement under uncertainty. *Psychological Science*, 14:195–200, 2003.

[10] D. Fudenberg and J. Tirole. Perfect bayesian equilibrium and secuential equilibrium. *Journal of Economic Theory*, 53:236–260, 1991.

[11] J. Geanakoplos. Nash and walras equilibrium via brouwer. *Economic Theory*, 21:585–603, 2002.

[12] J.C. Harsanyi and R. Selten. *A General Theory of Equilibrium Selection in Games.* The MIT Press, 1988.

[13] W. Haskel and F. Ordóñez. Mixed integer optimization methods for solving stackelberg security games. Sin publicar, 2015.

[14] M. Jain, C. Kiekintveld, F. Ordóñez, and M. Tambe. Bayesian security games with arbitrary schedules. *Journal of Artificial Intelligence Research*, 2012.

[15] M. Jain, J. Tsai, J. Pita, C. Kiekintveld, S. Rathi, M. Tambe, and F. Ordóñez. Software assistants for randomized patrol planning for the lax airport police and the federal air marshal service. *Interfaces*, 40:267–290, 2010.

[16] D. M. Kreps. *A Course in Microeconomic Theory*. Princeton University Press, 1990.

[17] L. Mallozzi and J. Morgan. Mixed strategies for hierarchical zero-sum games. *Annals of the International Society of Dynamics Games*, 6:65–77, 2002.

[18] J. Nash. Noncooperative games. *Annals of Mathematics*, 54:286–295, 1951.

[19] F. Ordóñez and N. E. Stier-Moses. Robust wardrop equilibrium. *NET-COOP*, pages 247–256, 2007.

[20] M. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

[21] M. J. Osborne. *An Introduction to Game Theory*. Oxford University Press, 2000.

[22] P. Paruchuri, J. P. Pearce, M. Tambe, F. Ordóñez, and S. Kraus. An efficient heuristic approach for security against multiple adversaries. *In International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2007.

[23] J. Pita, M. Jain, F. Ordóñez, M. Tambe, S. Kraus, and R. Magori-Cohen. Robustness in stackelberg games: Towards addressing rationality and observational uncertainties. *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, 2009.

[24] M. F. Sain. Cambio institucional y nueva policía en argentina. *V Encontro do Fórum Brasileiro de Seguranca Pública*, 2011.

[25] Y. Shoham, R. Powers, and T. Grenager. If multi-agent learning is the answer, what is the question? *AIJ*, 171(7):365–377, 2007.

[26] M. Tambe. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press, 2012.

[27] J. Tsai, J. Pita, C. Kiekintveld, M. Tambe, and F. Ordóñez. Computing optimal randomized resource allocations for massive security games. *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 689–696, 2009.

[28] J. von Neumann and O. Morgenstern. *The Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[29] H. F. von Stackelberg. *Market Structure and Equilibrium (original: Marktform und Gleichgewicht)*. Springer Berlin Heidelberg, 2010. Traducido por D. Bazin, R. Hill and L. Urch.

[30] B. von Stengel and S. Zamir. Leadership with commitment to mixed strategies. *CDAM Research Report LSE-CDAM-2004-01*, 2004.

[31] Z. Yin, M. Jain, M. Tambe, and F. Ordóñez. Risk-averse strategies for security games with execution and observational uncertainty. *Conference on Artificial Intelligence (AAAI)*, 2011.