CIMAT
Centro de Investigación en Matemáticas, A.C.

UN CONCEPTO DE EMPAREJAMIENTO ESTABLE CON EQUIDAD DE GÉNERO

T E S I S
Que para obtener el grado de
Maestro en Ciencias
con Orientación en
Matemáticas Básicas

Presenta
Fabio Enrique García Chica

Director de Tesis:
Dr. Francisco Sánchez Sánchez

____
Autorización de la versión final

Guanajuato, Gto., 20 de junio de 2018

A mis padres, el emparejamiento más estable que conozco

Agradecimientos

Al profesor Francisco Sánchez Sánchez, quien pacientemente guió mis esfuerzos a lo largo de los últimos tres semestres.

A los sinodales, profesores William Olvera y José Ignacio Barradas, por su amable disposición.

A quienes son y han sido mis colegas en el estudio de la teoría de juegos, en particular Miguel Vargas, David Benítez y Oliver Juárez, por sus contribuciones (balanceadas) a mis investigaciones para el trabajo.

A los docentes del CIMAT y a mis colegas estudiantes, una comunidad de pares que durante dos años he tenido el gusto de llamar hogar.

A mis padres, que han sido testigos de los deleites, preocupaciones, incertidumbres, frustraciones y satisfacciones que vienen con la elaboración de una tesis, y que nunca han dejado de alentarme.

Al CONACyT, que con su apoyo económico permitió mis estudios de maestría y la realización de este trabajo.

Contenido

Capítulo 1

Introducción

Dado un conjunto de hombres y mujeres que tienen diversas preferencias sobre el género opuesto, ¿cómo hacer parejas con ellos, de manera que no haya un hombre y una mujer que se prefieran mutuamente que a sus respectivas parejas asignadas? Éste es el problema del emparejamiento estable, un conocido problema con aplicaciones en la asignación de médicos a hospitales y de estudiantes a instituciones educativas [14, 19].

El problema del emparejamiento estable fue planteado y resuelto en 1962 por Gale y Shapley [1], engendrando luego un amplio campo de trabajo con diversas variantes, extensiones y aplicaciones [14], y con vínculos a la teoría de juegos [5, 21], a la combinatoria [4], al diseño de mercados [3, 23] y a la computación [2, 10].

La solución propuesta por Gale y Shapley es conocida por su asimetría de género: cuando existe más de un emparejamiento estable, el procedimiento siempre le da a uno de los géneros el mejor resultado que puede, y al otro género el peor resultado que puede. Esto ha motivado trabajos subsiguientes [13, 21, 24] en busca de una solución al problema del emparejamiento estable, de la cual pueda decirse que cuenta con equidad de género. Sin embargo, un resultado de Masarani y Gokturk [13] afirma que no es posible dar un método de solución que satisfaga varias nociones de justicia, entre ellas la equidad de género.

A la luz de ese resultado de imposibilidad, el presente trabajo es un intento por plantear un enfoque distinto para el problema de la equidad de género en emparejamientos estables. Para esto, se propondrá un concepto de emparejamiento que consiste en asignarle múltiples parejas a cada persona. La noción de estabilidad que se dará para estos emparejamientos multivaluados generaliza la noción de estabilidad para los emparejamientos del problema original, y se verá que una solución multivaluada siempre permite recuperar una o más soluciones del problema original. Además, se dará un procedimiento para obtener emparejamientos multivaluados estables que sí satisface las nociones de justicia dadas, y que en particular tiene equidad de género.

El trabajo consta de cinco capítulos, siendo el primero la presente introducción. En el capítulo 2 se presentan las definiciones y resultados básicos sobre emparejamientos estables y listas de preferencias; se plantea el algoritmo de Gale–Shapley para hallar emparejamientos estables, y se demuestran algunos resultados clásicos sobre la estructura del conjunto de emparejamientos estables para una lista de preferencias fija.

El capítulo 3 está dedicado a los trabajos previos sobre equidad de género, con énfasis en las nociones de justicia involucradas en el resultado de imposibilidad de [13]. Se describe también un enfoque que busca la equidad de género mediante la minimización de alguna función que se postula como medida de la inequidad de un emparejamiento [24], ilustrando con ejemplos cómo dichas funciones pueden no capturar adecuadamente la equidad buscada. Además, se comenta una variante del problema del matrimonio estable conocida como el problema de los compañeros de cuarto, en la cual se trabaja con un solo género en vez de dos. Si bien esta variante se ha estudiado por sí misma, el hecho de que el problema no distingue entre géneros hace que tenga sentido estudiarlo en el contexto de la equidad de género.

En el capítulo 4 se presentará el resultado principal: se definirán los emparejamientos multivaluados y se presentará la noción de estabilidad para ellos. Seguidamente, se describirá un algoritmo para reducir listas de preferencias, que funciona como paso intermedio en la formulación del algoritmo definitivo, y se demostrarán algunas de sus propiedades. A continuación se formulará el algoritmo de emparejamiento multivaluado y se demostrará que satisface equidad de género; finalmente, se ilustrará el funcionamiento del algoritmo mediante un ejemplo. El trabajo cierra con el capítulo 5 a modo de conclusión.

Capítulo 2

Nociones y resultados preliminares

En este capítulo inicial se dará una descripción del problema de los emparejamientos estables y se mostrarán los resultados básicos de la teoría.

2.1. Emparejamientos estables

Para empezar, se definirán los elementos que constituyen la materia prima del problema, a saber: la noción de emparejamiento entre dos conjuntos, la estructura de preferencias de los participantes y el concepto de estabilidad.

2.1.1. Emparejamientos y preferencias

Considérense dos conjuntos finitos disjuntos no vacíos H y M. En general, a los elementos de H se les llamará *hombres*, y a los de M, *mujeres*. Los elementos de $H \cup M$, indistintamente del género, se llamarán *individuos*, *participantes* o *personas*.

El interés es formar parejas de hombre y mujer, es decir, de un miembro de H y un miembro de M. No es requisito que H y M tengan el mismo número de elementos; en cualquier caso, es posible que algunos individuos queden sin pareja, y en tal caso se dirá que están *solteros*.

Definición 2.1. Un *emparejamiento entre* H y M es una función biyectiva

$$\mu : H \cup M \to H \cup M$$

que satisface lo siguiente:

- $\mu(\mu(x)) = x$ para todo $x \in H \cup M$;
- para todo $h \in H$, $\mu(h) \in M \cup \{h\}$ (μ empareja a cada hombre con una mujer o consigo mismo);
- para toda $m \in M$, $\mu(m) \in H \cup \{m\}$ (μ empareja a cada mujer con un hombre o consigo misma).

Si $\mu(i) = i$, se dice que i está *soltero* en μ.

Obsérvese que, por la primera condición, $\mu(i) = j$ si y sólo si $\mu(j) = i$. Por tanto, cuando $i \neq j$ puede decirse que i y j forman una *pareja*, o que son pareja, o que i es la pareja de j.

La razón por la cual no todos los emparejamientos son igualmente sensatos es que los individuos tienen preferencias sobre sus parejas potenciales: dados dos emparejamientos distintos μ_1 y μ_2, cada participante puede estar más satisfecho o menos con uno de ellos que con el otro.

Concretamente, a cada hombre $h \in H$ le corresponde un orden de preferencias sobre $M \cup \{h\}$. En adelante se supondrá que este orden es total; es decir, dadas dos alternativas (dos mujeres, o una mujer y la condición de quedarse soltero), h siempre prefiere una de las dos alternativas, y lo hace de manera coherente. En artículos anteriores [14, 15] se ha tratado la posibilidad de que los órdenes de preferencias no sean totales, concretamente, cuando un individuo puede ser indiferente entre dos alternativas; una manera sensata de abordar el problema es "desempatar" las indiferencias de manera arbitraria.

Estos órdenes de preferencias pueden definirse formalmente de muchas maneras, todas ellas equivalentes entre sí y todas ellas útiles para distintos propósitos técnicos. La que se empleará en este trabajo asigna a cada hombre $h \in H$ una relación de orden total estricto $\succ_h$ en $M \cup \{h\}$ tal que, para cada $i, j \in M \cup \{h\}$,

$$i \succ_h j \text{ si y sólo si } h \text{ prefiere a } i \text{ sobre } j.$$

La alternativa h se interpreta como la posibilidad de quedarse soltero: $h \succ_h m$ si y sólo si h prefiere estar soltero que tener a m como pareja.

Desde luego, no son sólo los hombres los que tienen preferencias con respecto al género opuesto: cada mujer $m \in M$ cuenta con su propio orden de preferencias $\succ_m$ sobre $H \cup \{m\}$.

Ejemplo. Supónganse las siguientes preferencias para tres hombres α, β, γ y tres mujeres A, B, C, donde cada columna corresponde a un individuo, y sus alternativas aparecen enumeradas de la más preferida a la menos preferida:

α	β	γ	A	B	C
A	A	C	α	β	γ
B	β	A	β	γ	C
C	B	γ	γ	B	α
α	C	B	A	α	β

Así pues, por ejemplo, en las preferencias de α,

$$A \succ_\alpha B \succ_\alpha C \succ_\alpha \alpha,$$

mientras que C tiene preferencias

$$\gamma \succ_C C \succ_C \alpha \succ_C \beta.$$

Se entiende que α aceptaría estar con cualquier mujer, mientras que C sólo aceptaría estar con γ y de lo contrario preferiría quedarse soltera.

Definición 2.2. Para cada individuo $i \in H \cup M$ sea $\mathcal{P}i$ un orden de preferencias. A la colección de todos estos órdenes de preferencias

$$\{(i, \mathcal{P}i) \mid i \in H \cup M\},$$

se la llamará un *perfil de preferencias entre* H y M, y se la denotará por $\mathcal{P}$. Esto es, en efecto, una función que a cada individuo i le asigna su orden de preferencias $\mathcal{P}i = \succ_i$.

Así pues, para un perfil de preferencias $\mathcal{P}$ entre H y M e individuos $i, j, k \in H \cup M$, los siguientes enunciados son maneras equivalentes de referirse al mismo hecho:

$$i \succ_k j \text{ en el perfil } \mathcal{P},$$
$$\text{en el perfil } \mathcal{P},\ k \text{ prefiere a } i \text{ sobre } j.$$

2.1.2. Estabilidad

En general, no será posible satisfacer perfectamente todas las preferencias de los participantes con un emparejamiento. Por ejemplo, si dos hombres h_1 y

h_2 coinciden en una misma mujer m como la que más les gusta, al menos uno de los dos no obtendrá a su pareja preferida, pues no es posible emparejar a m simultáneamente con ambos hombres.

Sin embargo, si m prefiere a h_1 sobre h_2, entonces un emparejamiento μ que asigne $\mu(m) = h_2$ no debe ser admitido, ya que m y h_1 siempre estarían dispuestos a abandonar a sus parejas para irse juntos:

- m prefiere a h_1 sobre su pareja asignada h_2, y
- h_1 prefiere a m que a su pareja asignada, sea cual sea, ya que m es su mujer preferida.

En este caso, la decisión de m y h_1 de irse juntos y abandonar a sus parejas iría en contra del emparejamiento μ, que es la razón por la cual μ se considera una solución inadecuada: hay personas que tienen ocasión de oponerse al emparejamiento. Lo que se espera de un emparejamiento, pues, es que esta clase de abandonos no sea posible: si alguien prefiere estar con una persona distinta de su pareja, esa persona ya está emparejada con alguien que considera mejor.

Definición 2.3. Sea $\mathcal{P}$ un perfil de preferencias entre H y M. Se dice que las personas i y j *bloquean* un emparejamiento μ (con respecto al perfil $\mathcal{P}$) si se cumple lo siguiente:

$$j \succ_i \mu(i) \qquad\qquad \text{y} \qquad\qquad i \succ_j \mu(j).$$

En otras palabras, i y j bloquean μ si se prefieren mutuamente que a las parejas que les asigna μ. Si existen personas que bloquean μ, se dice que μ es un emparejamiento *bloqueado* (con respecto a $\mathcal{P}$). Un emparejamiento que no es bloqueado se llama *estable* (con respecto a $\mathcal{P}$).

Nótese que no es necesario que i y j sean individuos distintos. Si en la definición se escoge $i = j$, las condiciones se reducen a que i prefiere estar soltero que ser pareja de $\mu(i)$.

Teorema 2.4. Un emparejamiento μ es estable con respecto a un perfil de preferencias dado si y sólo si, para todo $i, j \in H \cup M$, se cumple que

$$\text{si } j \succ_i \mu(i), \text{ entonces } \mu(j) \succ_j i.$$

Demostración. Si μ es estable y $j \succ_i \mu(i)$, entonces $\mu(j) \succeq_j i$, pues de lo contrario i y j bloquearían μ. Ahora bien, $\mu(i) \neq j$ ya que estrictamente $j \succ_i \mu(i)$. Se sigue que estrictamente $\mu(j) \succ_j i$.

Recíprocamente, si para todo $i, j \in H \cup M$ se cumple que $\mu(j) \succ_j i$ siempre que $j \succ_i \mu(i)$, entonces μ no tiene bloqueos y es por tanto estable. $\square$

Ejemplo. Considérese el siguiente perfil de preferencias entre tres hombres $H = \{\alpha, \beta, \gamma\}$ y tres mujeres $M = \{A, B, C\}$:

α	β	γ	A	B	C
A	B	A	β	β	γ
B	C	C	γ	α	α
C	A	B	α	γ	β
α	β	γ	A	B	C

El emparejamiento μ que asigna las parejas

$$\mu(\alpha) = C, \qquad \mu(\beta) = B, \qquad \mu(\gamma) = A,$$

es estable, como lo comprueba una verificación directa del teorema 2.4:

$$\begin{array}{rcl} A \succ_\alpha C = \mu(\alpha) & \text{pero} & \mu(A) = \beta \succ_A \alpha, \\ B \succ_\alpha C = \mu(\alpha) & \text{pero} & \mu(B) = \beta \succ_B \alpha, \\ \beta \succ_A \gamma = \mu(A) & \text{pero} & \mu(\beta) = B \succ_\beta A, \\ \gamma \succ_C \alpha = \mu(C) & \text{pero} & \mu(\gamma) = A \succ_\gamma C. \end{array}$$

En cambio, el emparejamiento ν que asigna

$$\mu(\alpha) = A, \qquad \mu(\beta) = B, \qquad \mu(\gamma) = C,$$

no es estable, pues en particular γ y A lo bloquean:

$$A \succ_\gamma C = \mu(\gamma) \qquad \text{y} \qquad \gamma \succ_A \alpha = \mu(A).$$

Definición 2.5. Sean $i, j \in H \cup M$. Si $j \succ_i i$, se dice que j es una *pareja admisible* de i. En cambio, si $i \succ_i j$, se dice que j es una *pareja descalificada* por i. Nótese que estas preferencias son estrictas; en particular, i no es pareja admisible ni descalificada de sí misma.

Dado que el interés del problema es hallar emparejamientos estables para un perfil de preferencias, es claro que los emparejamientos bloqueados no deben tenerse en cuenta. En la siguiente subsección se mostrará que sólo es necesario saber las preferencias de cada individuo con respecto a sus parejas admisibles; en consecuencia, las parejas descalificadas se omitirán al escribir las preferencias de cada persona.

2.1.3. Listas de preferencias

Las nociones de bloqueo y estabilidad dependen de los órdenes de preferencias de los participantes. En este trabajo, para enunciar algunos resultados se aludirá a varios perfiles de preferencias simultáneamente; cuando sea necesario aclarar que el bloqueo ocurre con respecto a un perfil de preferencias $\mathcal{P}$, se dirá que μ es un perfil $\mathcal{P}$-bloqueado, o que los individuos i y j $\mathcal{P}$-bloquean μ, según el caso. Asimismo, se hablará de emparejamientos $\mathcal{P}$-estables.

Una pregunta natural es cuándo pueden describirse los emparejamientos estables de $\mathcal{Q}$ a partir de los de $\mathcal{P}$ si se sabe dónde coinciden $\mathcal{P}$ y $\mathcal{Q}$. El siguiente resultado ofrece un caso en el cual esto es posible.

Teorema 2.6. Sean $\mathcal{P}$ y $\mathcal{Q}$ perfiles de preferencias entre H y M, μ un emparejamiento entre H y M. Sea $i \in H \cup M$, y supóngase que

a) los órdenes $\mathcal{P}k$ y $\mathcal{Q}k$ son iguales para todo $k \neq i$;

b) μ es $\mathcal{P}$-estable;

c) para todo k, si $\mu(i) \succ_i k$ en $\mathcal{P}$, entonces también $\mu(i) \succ_i k$ en $\mathcal{Q}$.

Entonces μ es $\mathcal{Q}$-estable.

Demostración. Sean k, ℓ tales que $\ell \succ_k \mu(k)$ en $\mathcal{Q}$. En virtud del teorema 2.4, basta ver que $\mu(\ell) \succ_\ell k$ en $\mathcal{Q}$. Para empezar, nótese que

$$\ell \succ_k \mu(k) \text{ en } \mathcal{P}.$$

Esto es cierto por la hipótesis a) cuando $k \neq i$, y cuando $k = i$ se sigue de la hipótesis c) y de la totalidad del orden. Usando ahora la $\mathcal{P}$-estabilidad de μ y el teorema 2.4,

$$\mu(\ell) \succ_\ell k \text{ en } \mathcal{P}.$$

De aquí se concluye que

$$\mu(\ell) \succ_\ell k \text{ en } \mathcal{Q},$$

lo cual es cierto cuando $\ell \neq i$ por la hipótesis a), y cuando $\ell = i$, por la hipótesis c). □

El siguiente resultado da una condición suficiente para que dos perfiles $\mathcal{P}$ y $\mathcal{Q}$ tengan los mismos emparejamientos estables.

Teorema 2.7. Sean $\mathcal{P}$ y $\mathcal{Q}$ perfiles de preferencias entre H y M tales que, para cada $i, j, k \in H \cup M$,

$$j \succeq_i k \succeq_i i \text{ en } \mathcal{P} \qquad \text{si y sólo si} \qquad j \succeq_i k \succeq_i i \text{ en } \mathcal{Q}$$

(es decir, cada individuo tiene las mismas parejas admisibles en ambos perfiles y las ordena de idéntica manera). Entonces un emparejamiento μ es $\mathcal{P}$-estable si y sólo si es $\mathcal{Q}$-estable.

Demostración. Por simetría, basta comprobar que si μ es $\mathcal{P}$-estable entonces también es $\mathcal{Q}$-estable. Obsérvese primero que, negando los enunciados de la hipótesis y poniendo $k = i$, se tiene

$$i \succ_i j \text{ en } \mathcal{P} \qquad \text{si y sólo si} \qquad i \succ_i j \text{ en } \mathcal{Q}.$$

Supóngase la estabilidad con respecto a $\mathcal{P}$. Sea $i_1 \in H \cup M$, y defínase un perfil de preferencias $\mathcal{P}^1$ tal que

$$\mathcal{P}^1 k = \mathcal{P} k \text{ para } k \neq i,$$
$$\mathcal{P}^1 i = \mathcal{Q} i,$$

es decir, $\mathcal{P}^1$ se obtiene a partir de $\mathcal{P}$ reemplazando las preferencias de i por sus preferencias en $\mathcal{Q}$. Es claro que $\mathcal{P}$ y $\mathcal{P}^1$ satisfacen las hipótesis a) y b) del teorema 2.6. Para concluir que μ es $\mathcal{P}^1$-estable, basta mostrar que se cumple la hipótesis c). En efecto, supóngase que $\mu(i) \succ_i k$ en $\mathcal{P}$. Considérense los siguientes casos:

- Si $k \succeq_i i$ en $\mathcal{P}$, entonces la hipótesis implica que
$$\mu(i) \succ_i k \succeq_i i \text{ en } \mathcal{P}^1.$$
- Si $i \succ_i k$ en $\mathcal{P}$, entonces $i \succ_i k$ en $\mathcal{P}^1$, como se ha visto. A su vez, $\mu(i) \succ_i i$ en $\mathcal{P}$ (de lo contrario, i bloquearía μ). Por la hipótesis,
$$\mu(i) \succ_i i \succ_i k \text{ en } \mathcal{Q}.$$

En ambos casos, $\mu(i) \succ_i k$ en $\mathcal{P}^1$, lo cual demuestra que μ es $\mathcal{P}^1$-estable.

Para terminar la prueba, constrúyanse perfiles

$$\mathcal{P}^2, \mathcal{P}^3, \ldots, \mathcal{P}^n$$

reemplazando sucesivamente las preferencias de cada individuo por sus preferencias en $\mathcal{Q}$ (donde $n = |H \cup M|$ y por consiguiente $\mathcal{P}^n = \mathcal{Q}$). El mismo razonamiento muestra que μ es $\mathcal{P}^r$-estable para todo $2 \leq r \leq n$, y por tanto μ es $\mathcal{Q}$-estable. $\square$

Este resultado implica que, al buscar un emparejamiento estable para un perfil de preferencias $\mathcal{P}$, no es necesario tener en cuenta las parejas descalificadas.

Definición 2.8. Sea $\mathbb{P}(H, M)$ el conjunto de todos los perfiles de preferencias entre H y M. En $\mathbb{P}(H, M)$ se define una relación de equivalencia $\sim$ como sigue: $\mathcal{P} \sim \mathcal{Q}$ si y sólo si, para todo $i, j, k \in H \cup M$,

$$j \succeq_i k \succeq_i i \text{ en } \mathcal{P} \qquad \text{si y sólo si} \qquad j \succeq_i k \succeq_i i \text{ en } \mathcal{Q}.$$

Una clase de equivalencia $[\mathcal{P}]$ se llama una *lista de preferencias* entre H y M.

Así pues, el teorema 2.7 dice que el conjunto de emparejamientos estables está bien definido para cada lista de preferencias; en adelante la clase de equivalencia de $\mathcal{P}$ se denotará por el mismo símbolo $\mathcal{P}$ sin lugar a confusiones. La lista de preferencias del ejemplo de la página 5 puede escribirse como sigue:

α	β	γ	A	B	C
A	A	C	α	β	γ
B		A	β	γ	
C			γ		

2.2. El algoritmo de Gale–Shapley

El *problema del emparejamiento estable*, pues, consiste en determinar si para un perfil de preferencias $\mathcal{P}$ dado existe un emparejamiento $\mathcal{P}$-estable, y en encontrar dicho emparejamiento. En esta sección se verá que la respuesta es afirmativa en todos los casos, dando para ello un procedimiento que siempre localiza un emparejamiento estable.

2.2.1. Formulación

Gale y Shapley [1] formulan un procedimiento que, a partir de cualquier perfil de preferencias $\mathcal{P}$, obtiene un emparejamiento estable μ. El procedimiento es iterativo; en cada iteración,

1. cada hombre le hace una *propuesta* a la mujer que más le gusta que no lo haya rechazado todavía;

2. cada mujer que tenga propuesta de algún hombre *rechaza* a todos sus pretendientes excepto al que más le gusta de ellos.

Si bien el procedimiento formulado en [1] está diseñado para el caso particular en que todas las preferencias son completas (es decir, en que todas las parejas son admisibles), el algoritmo puede extenderse sin problemas especificando qué ocurre con las parejas inadmisibles. Se entiende que si un hombre es rechazado por la mujer que menos le gusta (dentro de sus parejas admisibles), entonces se queda soltero. Asimismo, una mujer siempre rechaza a un hombre que considera inadmisible, incluso si eso la lleva a rechazar todas sus propuestas.

El procedimiento termina en cuanto no hay rechazos, es decir, cuando ningún hombre le propone a una mujer que lo considera inadmisible y no hay dos hombres que le propongan a la misma mujer. Entonces, cada hombre que no se haya quedado soltero es emparejado con la mujer que más le gusta que no lo rechazó, y las mujeres que no hayan recibido propuestas admisibles se quedan solteras. Así pues, el algoritmo puede sintetizarse como sigue: *hombres proponen, mujeres rechazan.*

Es claro que el algoritmo termina en un número finito de pasos, ya que en cada iteración hay al menos un hombre que tiene estrictamente menos parejas potenciales que antes, a menos que no haya rechazos, en cuyo caso el algoritmo termina. También es claro que el algoritmo produce un emparejamiento, pues al final de cada iteración cada individuo tiene a lo más una pareja.

Ejemplo. Considérese la siguiente lista de preferencias entre cuatro hombres α, β, γ, δ y tres mujeres A, B, C:

α	β	γ	δ	A	B	C
A	B	B	C	δ	δ	β
C	C	A	A	γ	α	α
B		C	B	β	β	δ
					γ	γ

Obsérvese que las preferencias no son completas: β ha descalificado a A; y a su vez A descalifica por anticipado a α.

En la primera etapa del algoritmo, α hace una propuesta a A; β y γ proponen a B, y δ a C. Puesto que A no tiene a α como pareja admisible, α es rechazado. A su vez, B tiene propuestas de dos hombres, así que rechaza a γ quedándose con β.

En la segunda etapa, α y γ proponen a su segunda mujer preferida, respectivamente C y A. Ahora C tiene propuestas de α y δ, entre los cuales su preferido es α, así que δ es rechazado.

En la tercera etapa, δ propone a su siguiente mujer preferida, que es A. Así pues, γ es nuevamente rechazado, ya que A prefiere la nueva propuesta de δ.

Por último, γ hace una propuesta a su última pareja admisible, que es C, y es rechazado una vez más, pues C prefiere la propuesta que ya tiene de α.

Con eso termina el algoritmo, ya que ninguna mujer tiene más de una propuesta. El emparejamiento μ así obtenido es el siguiente:

$$\mu(\alpha) = C, \qquad \mu(\beta) = B, \qquad \mu(\delta) = A, \qquad \mu(\gamma) = \gamma.$$

Obsérvese que μ es un emparejamiento estable: si un hombre prefiere a una mujer sobre su pareja, esa mujer ya tiene una pareja mejor. Por ejemplo, δ preferiría estar con C, pero C ya está emparejada con α, a quien ella prefiere.

2.2.2. Estabilidad del resultado

Se verá entonces que el algoritmo de Gale–Shapley cumple con su cometido: localizar un emparejamiento estable dado un perfil de preferencias.

Teorema 2.9. (Gale–Shapley [1]). Sea $\mathcal{P}$ un perfil de preferencias entre H y M, y μ el emparejamiento obtenido al aplicar el algoritmo de Gale–Shapley con las preferencias de $\mathcal{P}$. Entonces μ es $\mathcal{P}$-estable.

Demostración. Supóngase que $i, j \in H \cup M$ son tales que

$$j \succ_i \mu(i),$$

Considérese primero el caso en que $i \in H$. Entonces esto sólo puede ocurrir si i fue rechazado por j en el desarrollo del algoritmo (y en particular, $j \in M$). Eso implica que j obtuvo una propuesta de otro hombre $h \succ_j i$. Ahora bien, en el algoritmo una mujer nunca rechaza a un hombre a menos que obtenga una propuesta de un hombre que considera aún mejor. Por tanto,

$$\mu(j) \succ_j i,$$

Supóngase en cambio que $i \in M$. Si $j \in H$, entonces

$$\mu(j) \succ_j i,$$

pues lo contrario estaría en contradicción con el caso anterior. Si en cambio $j \in M$, entonces $i = j$, y se tiene

$$i \succ_i \mu(i),$$

lo cual es absurdo, ya que en el algoritmo una mujer siempre rechaza a una pareja descalificada.

Se concluye que μ satisface el criterio del teorema 2.4 en todos los casos, y por tanto es estable. □

2.2.3. Asimetría de género en el algoritmo

En el enunciado del algoritmo salta a la vista la asimetría en los papeles que desempeñan hombres y mujeres. Desde luego, es posible ejecutar el algoritmo intercambiando los papeles: *mujeres proponen, hombres rechazan.* Nada garantiza, sin embargo, que el resultado en ambos casos sea igual.

Ejemplo. Con las preferencias del ejemplo de la página 11, el algoritmo *mujeres proponen, hombres rechazan* se desarrollaría de esta manera:

A y B proponen a δ, C propone a β. Entonces δ, quien tiene dos propuestas, rechaza a B, ya que prefiere a A.

A continuación B propone a su segunda pareja preferida, que es α. Ahora ningún hombre tiene más de una propuesta, por lo cual no hay rechazos.

El emparejamiento hallado es, pues,

$$\alpha \text{ con } B, \qquad \beta \text{ con } C, \qquad \delta \text{ con } A, \qquad \gamma \text{ queda soltero.}$$

Hay varios hechos dignos de observar si se comparan los resultados de los dos algoritmos:

- Los dos emparejamientos no son iguales, ya que B y C han intercambiado parejas. Esto demuestra que un emparejamiento estable para un perfil de preferencias en general no es único.

- La persona soltera es la misma en ambos casos: γ. En el teorema 2.15 se demostrará que esto no es coincidencia: si una persona está soltera en algún emparejamiento estable, entonces está soltera en todo emparejamiento estable.

- Todos los hombres prefieren el resultado *hombres-proponen* sobre el *mujeres-proponen*, o son indiferentes entre ambos. Análogamente, todas las mujeres prefieren el resultado *mujeres-proponen* sobre el *hombres-proponen*, o son indiferentes entre ellos. A continuación se verá que esto tampoco es coincidencia.

Definición 2.10. Sea $\mathcal{P}$ un perfil de preferencias entre H y M. El *conjunto de emparejamientos $\mathcal{P}$-estables* se denota por

$$\mathrm{Est}(\mathcal{P}) = \{\mu : H \cup M \to H \cup M \mid \mu \text{ es } \mathcal{P}\text{-estable}\}.$$

Si existe $\mu \in \mathrm{Est}(\mathcal{P})$ tal que $\mu(i) = j$, se dice que j es *pareja $\mathcal{P}$-factible* de i. El *conjunto de parejas $\mathcal{P}$-factibles* de un individuo $i \in H \cup M$ es

$$\mathrm{Fact}(i, \mathcal{P}) = \{\mu(j) \mid \mu \in \mathrm{Est}(\mathcal{P})\}.$$

La *mejor pareja factible* de un individuo i (con respecto a $\mathcal{P}$) es

$$\mathrm{MPF}(i, \mathcal{P}) = \max_{\mathcal{P}i} \mathrm{Fact}(i, \mathcal{P}),$$

donde el máximo es con respecto al orden $\mathcal{P}i = \succ_i$. A su vez, la *peor pareja factible* de i (con respecto a $\mathcal{P}$) es

$$\mathrm{PPF}(i, \mathcal{P}) = \min_{\mathcal{P}i} \mathrm{Fact}(i, \mathcal{P}).$$

Teorema 2.11. (Gale–Shapley [1]) Sea $\mathcal{P}$ un perfil de preferenciase entre H y M, y sea μ_H el emparejamiento dado por el algoritmo de Gale–Shapley *hombres-proponen*. Entonces

$$\mu_H(h) = \mathrm{MPF}(h, \mathcal{P}) \text{ para todo } h \in H.$$

Demostración. Supóngase a modo de contradicción que no todos los hombres obtienen a su mejor pareja factible con μ_H. Entonces en el desarrollo del algoritmo hay un hombre h que es el primero en ser rechazado por su mejor pareja factible $m = \mathrm{MPF}(h, \mathcal{P})$. Si h fue rechazado por m, es porque un hombre $h' \succ_m h$ también le propuso a m. Como m es pareja factible de h, existe $\nu \in \mathrm{Est}(\mathcal{P})$ tal que $\nu(m) = h$. Por la estabilidad de ν, debe cumplirse

$$\nu(h') \succ_{h'} m.$$

Como h' llegó a proponerle a m en el algoritmo, esto significa que h' fue rechazado por $\nu(h')$ antes de que h fuera rechazado por m. Pero esto contradice que h es el primer hombre en ser rechazado por su mejor pareja factible. □

Así pues, μ_H es *hombre-óptimo*, es decir, le da a cada hombre su mejor pareja factible. En la siguiente sección se demostrará el complemento de este resultado: μ_H le da a cada mujer su peor pareja factible, y por tanto es *mujer-pésimo*. Para esto, se describirá la estructura del conjunto $\text{Est}(\mathcal{P})$ en términos de un orden parcial derivado de las preferencias conjuntas de hombres y mujeres.

2.3. El conjunto de emparejamientos estables

Los resultados de esta sección son atribuidos por Knuth [2] a John H. Conway.

2.3.1. Comparaciones entre emparejamientos

Teorema 2.12. Sea $\mathcal{P}$ un perfil de preferencias entre H y M, y sean $\mu_1, \mu_2 \in \text{Est}(\mathcal{P})$. Sean $h_1, h_2 \in H$ distintos. Entonces

$$\max_{h_1}\{\mu_1(h_1), \mu_2(h_1)\} \neq \max_{h_2}\{\mu_1(h_2), \mu_2(h_2)\}.$$

Demostración. Supóngase a modo de contradicción que

$$\max_{h_1}\{\mu_1(h_1), \mu_2(h_1)\} = m = \max_{h_2}\{\mu_1(h_2), \mu_2(h_2)\}.$$

Sin pérdida de generalidad, sean $\mu_1(m) = h_1$ y $\mu_2(m) = h_2$. Entonces

$$m \succ_{h_1} \mu_2(h_1), \qquad m \succ_{h_2} \mu_1(h_2).$$

Dado que μ_1 es estable, debe cumplirse

$$h_2 \prec_m \mu_1(m) = h_1.$$

Análogamente, dado que μ_2 es estable, debe cumplirse

$$h_1 \prec_m \mu_2(m) = h_2,$$

lo cual es una contradicción. $\square$

Esta proposición dice que si a los dos hombres se les da a escoger entre su respectiva pareja en μ_1 y su pareja en μ_2, los dos hombres no escogen a la misma persona. Para enfatizar este resultado, considérese la función

$$\mu_1 \vee_H \mu_2 : H \to H \cup M,$$
$$(\mu_1 \vee_H \mu_2)(h) = \max_{\mathcal{P}h}\{\mu_1(h), \mu_2(h)\}$$

(la imagen es en $H \cup M$ ya que si h está soltero en ambos emparejamientos entonces $(\mu_1 \vee_H \mu_2)(h) = h$). Entonces el resultado puede enunciarse como sigue: si $\mu_1, \mu_2 \in \mathrm{Est}(\mathcal{P})$, entonces $\mu_1 \vee_H \mu_2$ es inyectiva.

Nótese que esto permite extender $\mu_1 \vee_H \mu_2$ a $H \cup M$ escribiendo

$$(\mu_1 \vee_H \mu_2)(m) = \begin{cases} h & \text{si } m = (\mu_1 \vee \mu_2)(h) \text{ para algún } h \in H, \\ m & \text{en otro caso.} \end{cases}$$

y la función así extendida es un emparejamiento entre H y M. A continuación se verá que dicho emparejamiento es estable.

Teorema 2.13. Sean $\mu_1, \mu_2 \in \mathrm{Est}(\mathcal{P})$, $i_0 \in H \cup M$, $i_1 = \mu_1(i_0)$. Supóngase que $i_1 \succ_{i_0} \mu_2(i_0)$. Considérese la sucesión $\{i_n\} \subseteq H \cup M$ obtenida aplicando μ_1 y μ_2 alternadamente, empezando en i_0:

$$i_1 = \mu_1(i_0),$$
$$i_2 = \mu_2(\mu_1(i_0)) = \mu_2(i_1),$$
$$i_3 = \mu_1(\mu_2(\mu_1(i_0))) = \mu_1(i_2),$$
$$\vdots$$

Entonces

$$i_{n+1} \succ_{i_n} i_{n-1} \text{ para todo } n \geq 1.$$

Demostración. La prueba es por inducción. El caso $n = 1$ resulta por la estabilidad de μ_2: dado que $i_1 \succ_{i_0} \mu_2(i_0)$, se tiene $i_2 = \mu_2(i_1) \succ_{i_1} i_0$.

Supóngase la afirmación para cierto $n \geq 1$. Si n es impar, entonces $i_n = \mu_1(i_{n-1})$ e $i_{n+1} = \mu_2(i_n)$. Por tanto, la hipótesis de inducción afirma que

$$\mu_2(i_n) = i_{n+1} \succ_{i_n} i_{n-1} = \mu_1(i_n).$$

Por la estabilidad de μ_1,

$$i_{n+2} = \mu_1(i_{n+1}) \succ_{i_{n+1}} i_n,$$

lo cual es la afirmación para $n + 1$. La prueba cuando n es par es análoga intercambiando μ_1 y μ_2. $\square$

Teorema 2.14. Sean $\mu_1, \mu_2 \in \mathrm{Est}(\mathcal{P})$, $m \in M$. Entonces

$$(\mu_1 \vee_H \mu_2)(m) \in \{\mu_1(m), \mu_2(m)\}.$$

Demostración. Si $(\mu_1 \vee_H \mu_2)(m) = h \in H$, entonces $m = (\mu_1 \vee_H \mu_2)(h)$, de forma que $m = \mu_1(h)$ o bien $m = \mu_2(h)$.

Supóngase en cambio que $(\mu_1 \vee_H \mu_2)(m) = m$. Quiere verse que $\mu_1(m) = m$ o $\mu_2(m) = m$. Supóngase a modo de contradicción que $h_1 = \mu_1(m) \neq m$ y $h_2 = \mu_2(m) \neq m$. Entonces $h_1 \neq h_2$, pues de lo contrario

$$m = \mu_1(h_1) = \mu_2(h_1) = (\mu_1 \vee_H \mu_2)(h_1) \neq m.$$

Sin pérdida de generalidad, pues, $h_1 \succ_m h_2$. Considérese la sucesión $\{i_n\} \subseteq H \cup M$ obtenida aplicando μ_1 y μ_2 alternadamente como en el teorema 2.13, con $i_0 = m$, de forma que $i_1 = h_1$. Considérese ahora la composición $f = \mu_2 \circ \mu_1$, de forma que

$$i_{n+2} = f(i_n), \qquad \text{para } n \geq 0.$$

Como μ_1 y μ_2 son biyectivas, f también lo es. Por tanto, existe un $p > 0$ tal que

$$m = f^p(m) = i_{2p}.$$

Entonces $i_{2p-1} = \mu_2(m) = h_2$ e $i_{2p-2} = \mu_1(h_2)$. Por el teorema 2.13,

$$i_{2p} \succ_{i_{2p-1}} i_{2p-2},$$

es decir,

$$m \succ_{h_2} \mu_1(h_2).$$

Puesto que $m = \mu_2(h_2)$, esto implica que

$$m = \max\{\mu_1(h_2), \mu_2(h_2)\} = (\mu_1 \vee_H \mu_2)(h_n) \neq m,$$

lo cual es una contradicción. $\square$

Teorema 2.15. (Gale–Sotomayor [6]). Sea $\mu_0 \in \mathrm{Est}(\mathcal{P})$ y sea $m \in M$ tal que $\mu_0(m) = m$. Entonces $\mu(m) = m$ para todo $\mu \in \mathrm{Est}(\mathcal{P})$.

Demostración. Supóngase a modo de contradicción que $h = \mu(m) \neq m$. Entonces, por el teorema 2.14,

$$(\mu \vee_H \mu_0)(m) \in \{m, h\}.$$

Se procede entonces por casos:

- Si $(\mu \vee_H \mu_0)(m) = h$, entonces, por la definición de $\mu \vee_H \mu_0$,

$$m \succ_h \mu_0(h),$$

y por la estabilidad de μ_0,

$$m = \mu_0(m) \succ_m h = \mu(m),$$

contradiciendo la estabilidad de μ.

- Si $(\mu \vee_H \mu_0)(m) = m$, entonces

$$m \neq (\mu \vee_H \mu_0)(h) \succ_h m = \mu(h),$$

lo que también contradice la estabilidad de μ.

Se concluye que $\mu(m) = m$. □

Teorema 2.16. (Conway). Sean $\mu_1, \mu_2 \in \mathrm{Est}(\mathcal{P})$. Entonces

$$\mu_1 \vee_H \mu_2 \in \mathrm{Est}(\mathcal{P}).$$

Demostración. Sean $m \in M$, $h \in H \cup \{m\}$ tales que

$$h \succ_m (\mu_1 \vee_H \mu_2)(m).$$

Por el teorema 2.14, esto implica que

$$h \succ_m \mu_1(m) \qquad \text{o bien} \qquad h \succ_m \mu_2(m).$$

Como μ_1 y μ_2 son estables, esto equivale a que

$$\mu_1(h) \succ_h m \qquad \text{o bien} \qquad \mu_2(h) \succ_h m.$$

Por la definición de $\mu_1 \vee_H \mu_2$,

$$(\mu_1 \vee_H \mu_2)(h) \succ_h m.$$

Esto demuestra que m no hace parte de ningún bloqueo, es decir, un bloqueo de $\mu_1 \vee_H \mu_2$ no puede contener mujeres. Por tanto, si dos individuos i y j bloquean $\mu_1 \vee_H \mu_2$, entonces $i = j \in H$. Esto significa que

$$h \succ_h (\mu_1 \vee_H \mu_2)(h) \succeq_h \mu_1(h),$$

contradiciendo la estabilidad de μ_1. Se concluye que μ no tiene bloqueos, y por tanto es estable. □

2.3.2. El retículo de emparejamientos estables

El teorema 2.16 define una operación

$$\vee_H : \mathrm{Est}(\mathcal{P}) \times \mathrm{Est}(\mathcal{P}) \to \mathrm{Est}(\mathcal{P}).$$

Este teorema, como los resultados de la subsección anterior, son igualmente válidos intercambiando hombres con mujeres y definiendo la operación análoga

$$\vee_M : \mathrm{Est}(\mathcal{P}) \times \mathrm{Est}(\mathcal{P}) \to \mathrm{Est}(\mathcal{P})$$

que le da a las mujeres a escoger la pareja que más les gusta entre dos emparejamientos dados. En particular, el análogo del teorema 2.15 se tiene también para cada $h \in H$ usando la operación $\vee_M$ en lugar de $\vee_H$. En consecuencia, puede hablarse sin ambigüedades de los *solteros de un perfil de preferencias* $\mathcal{P}$, a saber, las personas que se quedan solteras en cualquier emparejamiento estable. Este resultado será útil más adelante.

A continuación se dará la relación entre las dos operaciones $\vee_H$ y $\vee_M$. Se empieza observando que $\vee_H$ y $\vee_M$ son asociativas, lo cual se sigue de la definición y de la asociatividad del máximo con respecto a un orden lineal: para $h \in H$ y $\mu_1, \mu_2, \mu_3 \in \mathrm{Est}(\mathcal{P})$,

$$\max_{\mathcal{P}h}\{\mu_1(h), \max_{\mathcal{P}h}\{\mu_2(h), \mu_3(h)\}\} = \max_{\mathcal{P}h}\{\max_{\mathcal{P}h}\{\mu_1(h), \mu_2(h)\}, \mu_3(h)\}.$$

Teorema 2.17. Defínase una relación $\geq_H$ en $\mathrm{Est}(\mathcal{P})$ por

$$\mu_1 \geq_H \mu_2 \text{ si y sólo si } \mu_1 \vee_H \mu_2, \qquad \text{para todo } \mu_1, \mu_2 \in \mathrm{Est}(\mathcal{P}).$$

Entonces $\geq_H$ es un orden parcial en $\mathrm{Est}(\mathcal{P})$.

Demostración. $\geq_H$ es reflexiva pues $\mu_1 \vee_H \mu_1 = \mu_1$; es antisimétrica pues si $\mu_1 \geq_H \mu_2$ y $\mu_2 \geq_H \mu_1$, entonces

$$\mu_1 = \mu_1 \vee_H \mu_2 = \mu_2;$$

y es transitiva pues si $\mu_1 \geq_H \mu_2$ y $\mu_2 \geq_H \mu_3$, entonces

$$\mu_1 \vee_H \mu_3 = (\mu_1 \vee_H \mu_2) \vee_H \mu_3 = \mu_1 \vee_H (\mu_2 \vee_H \mu_3) = \mu_1 \vee_H \mu_2 = \mu_1,$$

lo que demuestra que $\mu_1 \geq_H \mu_3$. □

El enunciado $\mu_1 \geq_H \mu_2$ quiere decir que $\mu_1(h) \succeq_h \mu_2(h)$ para todo hombre h. Se tiene un orden análogo $\geq_M$ definido a partir de $\vee_M$.

Teorema 2.18. Sean $\mu_1, \mu_2 \in \mathrm{Est}(\mathcal{P})$. Entonces

$$\mu_1 \geq_H \mu_2 \qquad \text{si y sólo si} \qquad \mu_1 \leq_M \mu_2.$$

Demostración. Por simetría basta comprobar que si $\mu_1 \geq_H \mu_2$ entonces $\mu_1 \leq_M \mu_2$. Supóngase a modo de contradicción que $\mu_1 \nleq_M \mu_2$. Entonces existe al menos una mujer $m \in M$ tal que

$$\mu_1(m) \succ_m \mu_2(m).$$

Sea $i = \mu_1(m)$. Por la estabilidad de μ_2,

$$\mu_2(i) \succ_i m = \mu_1(i),$$

lo que contradice que $(\mu_1 \vee_H \mu_2)(i) = \mu_1(i)$. □

Esta proposición muestra que, en el conjunto de los emparejamientos estables, lo que beneficia a los hombres perjudica a las mujeres; en particular, se tiene el siguiente resultado:

Teorema 2.19. Sea μ_H el emparejamiento hombre-óptimo para el perfil de preferencias $\mathcal{P}$ entre H y M. Entonces

$$\mu_H = \bigvee_{\mu \in \mathrm{Est}(\mathcal{P})} \mu = \max_{\geq_H} \mathrm{Est}(\mathcal{P}) = \min_{\geq_M} \mathrm{Est}(\mathcal{P}).$$

Demostración. La primera igualdad se sigue de la asociatividad de $\vee_H$ y el teorema 2.11; la segunda, del teorema 2.17; y la tercera, del teorema 2.18. □

Así pues, el algoritmo de Gale–Shapley le da a las mujeres su peor pareja factible: la pareja que menos les gusta entre todas las que le son asignadas por emparejamientos estables.

Capítulo 3

Equidad de género

Como se vio en el capítulo anterior, al escoger un emparejamiento estable con respecto a un perfil de preferencias aparece una disyuntiva de género: lo que beneficia a los hombres perjudica a las mujeres. ¿Cómo escoger entonces un emparejamiento neutro, que no favorezca a ninguno de los dos géneros por encima del otro? En este capítulo se describirá cómo se ha tratado este problema en la literatura sobre emparejamientos.

3.1. Conceptos de justicia

Si para un perfil de preferencias dado $\mathcal{P}$ existe un único emparejamiento estable μ, cualquier método que se use para emparejar a los participantes debe dar con μ si es que su objetivo es asegurar la estabilidad. Por tanto, incluso si el resultado parece favorecer a un género mucho más que al otro, no puede decirse que se esté cometiendo una injusticia.

En cambio, si existe más de un emparejamiento $\mathcal{P}$-estable, entonces un método de emparejamiento debe elegir uno solo de ellos. Conviene asegurarse de que dicha elección se haga de manera justa y no de manera discriminatoria. En esta sección se formaliza la idea de un método de emparejamiento y se presentan diversas nociones de justicia que un tal método puede o no puede cumplir.

3.1.1. Indiferencia de género

Por razones técnicas, en la siguiente definición es necesario dar un dominio bien definido para la función, por lo cual se supone que existe un "conjunto de todas las personas", denotado aquí por X. Este conjunto puede ser infinito, pero es siempre no vacío.

Definición 3.1. Sea X un conjunto. Un *algoritmo de emparejamiento* es una función $\mathtt{A}$ que recibe como variables dos conjuntos finitos no vacíos disjuntos $H, M \subseteq X$ y un perfil de preferencias $\mathcal{P}$ entre H y M, y que produce como valor un emparejamiento $\mu = \mathtt{A}(H, M, \mathcal{P})$. Si $\mathtt{A}(H, M, \mathcal{P})$ es $\mathcal{P}$-estable para todo $H, M, \mathcal{P}$, se dice que $\mathtt{A}$ es un algoritmo de emparejamiento *estable*.

Definición 3.2. Un algoritmo $\mathtt{A}$ satisface el principio de *indiferencia de género* si $\mathtt{A}(H, M, \mathcal{P}) = \mathtt{A}(M, H, \mathcal{P})$ para cualesquiera conjuntos H, M y cualquier perfil de preferencias $\mathcal{P}$ entre H y M.

En esencia, un algoritmo es indiferente con respecto al género si su desarrollo no precisa tratar a los dos géneros de maneras distintas, como lo hace el algoritmo de Gale–Shapley. Los participantes deberían desear un algoritmo que satisfaga indiferencia de género, pues de lo contrario no sabrían a priori si su género será aquel que el algoritmo trate de manera preferente o desventajosa [13].

3.1.2. Indiferencia entre pares

Definición 3.3. Sean H' y M' conjuntos finitos disjuntos no vacíos, y sea

$$\theta : H \cup M \to H' \cup M'$$

una biyección tal que $\theta(H) = H'$ y $\theta(M) = M'$. Se dice que θ es una *permutación* de H y M en H' y M'.

Sea $\mathcal{P}$ un perfil de preferencias entre H y M. El *perfil permutado* $\theta\mathcal{P}$ es el perfil de preferencias entre H' y M' tal que, para todo $i, j, k \in H \cup M$,

$$\theta(i) \succ_{\theta(k)} \theta(j) \text{ en } \theta\mathcal{P} \qquad \text{si y sólo si } i \succ_k j \text{ en } \mathcal{P}.$$

En efecto, $\theta\mathcal{P}$ es el perfil obtenido a partir de $\mathcal{P}$ reemplazando cada aparición de i por $\theta(i)$ para todo $i \in H \cup M$.

Análogamente, si μ es un emparejamiento entre H y M, se define el *emparejamiento permutado* $\theta\mu$ entre H' y M' por

$$\theta\mu(\theta(i)) = \theta(\mu(i)), \text{ para todo } i \in H \cup M.$$

De este modo, $\theta\mu$ es el emparejamiento obtenido a partir de μ reemplazando a i por $\theta(i)$ para cada $i \in H \cup M$.

Ejemplo. Considérese la siguiente lista de preferencias $\mathcal{P}$ entre $H = \{\alpha, \beta\}$ y $M = \{A, B\}$:

α	β	A	B
A	B	β	α
B	A	α	β

Obsérvese que hay dos emparejamientos estables distintos: uno μ_H que da a cada hombre su pareja preferida y otro μ_M que da a cada mujer su pareja preferida. Considérese la permutación θ de H y M en M y H dada por

$$\theta(\alpha) = A, \qquad \theta(\beta) = B, \qquad \theta(A) = \beta, \qquad \theta(B) = \alpha.$$

Para obtener $\theta\mathcal{P}$ basta reemplazar cada aparición de i en la tabla por $\theta(i)$ para todo $i \in H \cup M$:

A	B	β	α
β	α	B	A
α	β	A	B

Se observa que $\theta\mathcal{P} = \mathcal{P}$. Además, $\theta\mu_H = \mu_M$ y $\theta\mu_M = \mu_H$, pues en cada caso es el género contrario el que obtiene sus parejas preferidas.

Definición 3.4. Un algoritmo $\mathtt{A}$ satisface el principio de *indiferencia entre pares* si para todo $H, M, \mathcal{P}$ y para cualquier permutación θ de H y M se cumple

$$\mathtt{A}(\theta(H), \theta(M), \theta\mathcal{P}) = \theta\mathtt{A}(H, M, \mathcal{P}).$$

En esencia, un algoritmo satisface indiferencia entre pares si no requiere en su desarrollo más información sobre los participantes que la estructura colectiva de sus órdenes de preferencias. Esto excluye los algoritmos de emparejamiento que en algún punto requieran escoger un individuo “de manera arbitraria”, “al azar”, ”en orden alfabético” o con respecto a otro criterio que no haga parte de los órdenes de preferencias.

Como con la indiferencia de género, cada participante debería preferir un algoritmo que satisfaga indiferencia entre pares. Lo contrario implica someterse a la posibilidad de que el criterio empleado para distinguir individuos le dé una desventaja en el emparejamiento [13].

3.1.3. Una dificultad con los conceptos de justicia

Masarani y Gokturk [13] hacen la siguiente observación:

Teorema 3.5. Ningún algoritmo de emparejamiento estable satisface al mismo tiempo indiferencia de género e indiferencia entre pares.

Demostración. Supóngase que $\mathtt{A}$ es un algoritmo de emparejamiento estable que satisface indiferencia de género e indiferencia entre pares. Considérense el perfil de preferencias $\mathcal{P}$ y la permutación θ de la página 23, que tiene dos emparejamientos estables μ_H y μ_M entre dos hombres y dos mujeres. De este modo,

$$\theta(H) = M, \qquad \theta(M) = H,$$
$$\theta\mathcal{P} = \mathcal{P},$$
$$\theta\mu_H = \mu_M \neq \mu_H, \qquad \theta\mu_M = \mu_H \neq \mu_M.$$

Por indiferencia entre pares,

$$\theta\mathtt{A}(H, M, \mathcal{P}) = \mathtt{A}(\theta(H), \theta(M), \theta\mathcal{P}) = \mathtt{A}(M, H, \mathcal{P}),$$

luego, por indiferencia de género,

$$\theta\mathtt{A}(H, M, \mathcal{P}) = \mathtt{A}(M, H, \mathcal{P}) = \mathtt{A}(H, M, \mathcal{P}),$$

lo cual es imposible ya que $\theta\mu \neq \mu$ para los dos emparejamientos estables. $\square$

La consecuencia de este resultado es que no es posible resolver el problema del emparejamiento estable de una manera que sea a la vez indistinta del género e indistinta del orden en que se consideran los participantes. El mismo contraejemplo muestra que es falsa una proposición de [21], que ofrece un método para convertir cualquier algoritmo de emparejamiento estable en uno con indiferencia de género. La demostración dada en el artículo falla precisamente cuando las preferencias de hombres y mujeres son isomorfas, es decir, cuando existe una permutación θ que envía H en M y M en H, tal que $\theta\mathcal{P} = \mathcal{P}$.

3.2. Medidas de inequidad

Vista la imposibilidad de resolver el problema de la indiferencia de género sin caer en la discriminación entre pares, resta preguntarse qué alternativas quedan. Sabiendo que los órdenes de preferencias son una medida de cuán "satisfecho" está un individuo con respecto a su pareja asignada, puede plantearse el objetivo de hallar un emparejamiento estable que deje a los participantes lo más "satisfechos" posible, independientemente del género. En esta sección se ofrecen distintas maneras de formalizar este propósito. A lo largo de esta sección se empleará la siguiente notación común:

Definición 3.6. Sea $\mathcal{P}$ un perfil de preferencias entre H y M. Se define el *conjunto de no solteros* $T \subseteq H \cup M$ por

$$T = \{i \in H \cup M \mid \mu(i) \neq i \text{ para cualquier } \mu \in \mathrm{Est}(\mathcal{P})\}.$$

(T está bien definido por el teorema 2.15). Para $i \in T$ y $\mu \in \mathrm{Est}(\mathcal{P})$, se define el *rango* de i en μ por

$$\#\mu(i) = |\{j \in H \cup M \mid j \succeq_i \mu(i)\}|,$$

efectivamente el lugar que ocupa $\mu(i)$ en la lista de preferencias de i.

Es claro que, cuanto más pequeño sea $\#\mu(i)$, más satisfecho estará i con su pareja en μ.

3.2.1. El costo de arrepentimiento

Una forma intuitiva de saber cuán inequitativo es un emparejamiento μ es hallar el valor máximo de $\#\mu(i)$ para $i \in H \cup M$. Esto consiste en determinar cuál de todos los participantes está "peor emparejado" y qué lugar ocupa su pareja en su orden de preferencias.

Definición 3.7. El *costo de arrepentimiento* [24] de μ se define por

$$r(\mu) = \max_{i \in T} \#\mu(i).$$

En general, al comparar dos emparejamientos μ_1 y μ_2, si $r(\mu_1) < r(\mu_2)$ entonces cabe esperar que μ_1 sea de alguna manera más equitativo que μ_2. Por tando, puede considerarse deseable hallar un emparejamiento μ que haga mínimo $r(\mu)$. Sin embargo, éste puede no ser el caso: considérese la siguiente lista de preferencias para cinco hombres y cinco mujeres.

α	β	γ	δ	ε	A	B	C	D	E
B	B	C	D	E	γ	δ	ε	γ	β
C	A	B	C	D	β	γ	δ	ε	α
D	E	A	B	C	ε	ε	β	β	δ
A	D	D	E	B	α	β	γ	δ	ε
E	C	E	A	A	δ	α	α	α	γ

Los únicos dos emparejamientos estables son un μ_1 que empareja

$$\mu_1(\alpha) = A, \quad \mu_1(\beta) = B, \quad \mu_1(\gamma) = C, \quad \mu_1(\delta) = D, \quad \mu_1(\varepsilon) = E,$$

y un μ_2 que empareja

$$\mu_2(\alpha) = E, \quad \mu_2(\beta) = A, \quad \mu_2(\gamma) = B, \quad \mu_2(\delta) = C, \quad \mu_2(\varepsilon) = D.$$

Puede verse que $r(\mu_1) = 4$ y $r(\mu_2) = 5$. Sin embargo,

- $\#\mu_2(i) = 2$ para todo $i \neq \alpha$;
- en cambio, $\#\mu_1(h) = 1$ para $h \in H \setminus \{\alpha\}$, pero $\#\mu_1(i) = 4$ para $i \in M \cup \{\alpha\}$.

En otras palabras, mejorar la situación de α perjudica gravemente a las mujeres en beneficio de los hombres. Si bien al minimizar el costo de arrepentimiento se consigue un resultado mejor para la persona peor emparejada, esto puede inducir más inequidad en los resultados para otras personas.

Dado un perfil de preferencias $\mathcal{P}$, el costo de arrepentimiento puede minimizarse sobre $\mathrm{Est}(\mathcal{P})$ con un algoritmo sencillo aunque ineficiente computacionalmente, basado en el algoritmo de Gale–Shapley. Para cada entero n desde 1 hasta $\max\{|H|, |M|\}$:

1. Sea $\mathcal{P}_n$ la lista de preferencias obtenida a partir de $\mathcal{P}$ dejando únicamente las primeras n preferencias de la lista de cada persona;
2. sea μ_n el emparejamiento de Gale–Shapley hombre-óptimo para $\mathcal{P}_n$;
3. comprobar si μ_n es $\mathcal{P}$-estable (esto puede hacerse de manera directa evaluando todas las posibles parejas bloqueadoras);
4. si μ_n es $\mathcal{P}$-estable, entonces $n = r(\mu_n)$ es mínimo; de lo contrario, pasar al siguiente valor de n.

Una aplicación repetida del teorema 2.6 muestra que un emparejamiento $\mathcal{P}_n$-estable debe ser $\mathcal{P}$-estable a menos que $\mu_n(i) = i$ para algún i que no es soltero en $\mathcal{P}$; es decir, si al reducir las listas de preferencias quedan solteras algunas personas que no lo estaban. Siendo éste el caso, se sigue que ningún emparejamiento $\mathcal{P}_n$-estable puede ser $\mathcal{P}$-estable, por el teorema 2.15. Por la construcción de $\mathcal{P}_n$, ningún emparejamiento $\mathcal{P}$-estable μ puede tener

$r(\mu) = n$. Esto implica que el algoritmo efectivamente minimiza el costo de arrepentimiento.

Gusfield [10] ofrece un algoritmo más eficiente para hallar un emparejamiento estable μ que minimiza $r(\mu)$, usando para ello la estructura de retículo de $\mathrm{Est}(\mathcal{P})$.

3.2.2. El costo igualitario

Una segunda idea para medir la inequidad de un emparejamiento $\mu \in \mathrm{Est}(\mathcal{P})$ a partir de las preferencias de $\mathcal{P}$ consiste en sencillamente sumar todos los valores $\#\mu(i)$ para $i \in H \cup M$. Al tener en cuenta los resultados de todos los individuos, esto previene situaciones como la del ejemplo anterior, en las cuales un solo participante perjudica a muchos.

Definición 3.8. El *costo igualitario* [24] de μ se define por

$$c(\mu) = \sum_{i \in T} \#\mu(i).$$

El inconveniente de esta idea es que tras la equidad general puede esconderse una inequidad de género mayor. El siguiente perfil de preferencias

α	β	γ	δ	A	B	C	D
A	B	C	D	β	γ	α	δ
D	D	A	A	δ	δ	β	α
B	C	D	B	γ	α	δ	β
C	A	B	C	α	β	γ	γ

tiene tres emparejamientos estables: los emparejamientos μ_H y μ_M que le dan a cada hombre (respectivamente, a cada mujer) su primera opción, y un emparejamiento μ_3 dado por

$$\mu_3(\alpha) = B, \qquad \mu_3(\beta) = C, \qquad \mu_3(\gamma) = A, \qquad \mu_3(\delta) = D.$$

Puede comprobarse que

$$c(\mu_H) = c(\mu_M) = 17, \qquad\qquad c(\mu_3) = 18,$$

a pesar de que μ_3 es claramente el más equitativo de los tres emparejamientos.

Irving, Leather y Gusfield [11] dieron el primer algoritmo para encontrar un emparejamiento estable μ que minimiza $c(\mu)$. El método consiste en explotar las propiedades combinatorias del conjunto de emparejamientos estables para reducir el problema a encontrar un subconjunto de peso máximo de una gráfica dirigida ponderada.

3.2.3. El costo de igualdad de sexos

La última de las tres medidas de inequidad que se examinan en esta sección es la única de ellas que se enfoca concretamente en las diferencias de género. La idea es calcular lo bien o mal emparejados que están en general los hombres, y compararlo con lo bien o mal emparejadas que están en general las mujeres.

Definición 3.9. El *costo de igualdad de sexos* [24] de μ se define por

$$d(\mu) = \left| \sum_{h \in H \cap T} \#\mu(h) - \sum_{m \in M \cap T} \#\mu(m) \right|.$$

Si bien al minimizar $d(\mu)$ se obtienen resultados más equitativos entre géneros, esto puede ocurrir a costa del bienestar general. Considérese esta lista de preferencias:

α	β	γ	δ	A	B	C	D
A	B	C	D	δ	δ	δ	δ
B	C	A	A	β	γ	α	α
D	A	B	B	α	β	γ	β
C	D	D	C	γ	α	β	γ

Los dos emparejamientos estables para estas preferencias son el emparejamiento μ_H que le da a cada hombre su pareja preferida, y otro μ_M que asigna

$$\mu_M(\alpha) = C, \qquad \mu_M(\beta) = A, \qquad \mu_M(\gamma) = B, \qquad \mu_M(\delta) = D.$$

Puede calcularse

$$d(\mu_H) = 6, \qquad\qquad d(\mu_M) = 4.$$

Sin embargo, puede argumentarse que μ_H es en general un resultado más deseable: en μ_H los hombres y D obtienen su primera opción, mientras que las demás mujeres obtienen su tercera opción. Pasar a μ_M hace que todos los participantes reciban su segunda, tercera o incluso cuarta alternativa, con excepción de D y δ. Esto se refleja en los costos igualitarios de los dos emparejamientos, que son respectivamente 14 y 18.

A diferencia de $r(\mu)$ y $c(\mu)$, no existen algoritmos eficientes para minimizar $d(\mu)$ [24]. La solución ineficiente consiste en enumerar todos los emparejamientos estables, calcular su costo de igualdad de sexos y escoger uno cuyo costo sea mínimo. Existen, sin embargo, algoritmos eficientes para acotar el valor de $d(\mu)$ [24].

3.2.4. Un comentario sobre medidas de inequidad

Como se ha visto, ninguna de las tres medidas de inequidad estudiadas es suficiente por sí sola para garantizar una verdadera equidad en el resultado. Puede argumentarse, sin embargo, a favor de un índice que combine las tres medidas de inequidad: en cada uno de los ejemplos de esta sección, calcular las tres medidas para los emparejamientos estables muestra que dos de ellas sí apuntan hacia el emparejamiento más equitativo en cada caso.

Hay, a pesar de todo, dos razones por las cuales puede considerarse que el enfoque basado en medidas de inequidad no es adecuado para enfrentar el problema de la equidad de género. La primera de ellas es que esto fuerza a asignarle un valor numérico a cada posición de la lista de preferencias de un individuo. Si

$$A \succ_h B \succ_h C,$$

entonces asignarle los valores $1, 2, 3$ a A, B, C no refleja de la misma manera las preferencias de h si h considera que B es casi tan preferible como A que si la considera apenas un poco más preferible que C. En consideración a esto, puede generalizarse el problema permitiendo que los individuos asignen pesos a cada posible pareja de una manera que refleje sus preferencias; este problema también ha sido estudiado [22].

La segunda razón por la cual este enfoque no es enteramente adecuado es que, incluso al obtener un algoritmo que minimiza alguna de las medidas de inequidad, el teorema 3.5 también se aplica a dicho algoritmo. Concretamente, cuando el mínimo se realiza en más de un emparejamiento estable, es imposible escoger entre ellos sin incurrir en discriminación de género o en discriminación entre pares.

3.3. El problema de los compañeros de cuarto

Una variante del problema del emparejamiento estable que conviene tener en cuenta a la hora de discutir asuntos de equidad de género es el problema de los compañeros de cuarto. En esta variante, los conjuntos H y M se ven reemplazados por un único conjunto I de $2n$ individuos, que tienen preferencias unos sobre otros, y el objetivo es formar parejas de individuos que van a habitar n cuartos con capacidad para dos personas. En consecuencia, en esta variante no está permitido que haya individuos solteros.

3.3.1. Planteamiento del problema

Definición 3.10. Sea I un conjunto finito no vacío con un número par de elementos. Un *emparejamiento* sobre I es una biyección $\mu : I \to I$ sin puntos fijos tal que $\mu^2 = \mu$. Un *perfil de preferencias* $\mathcal{P}$ es una función que a cada $i \in I$ le asigna un orden de preferencias $\mathcal{P}i = \succeq_i$ sobre I. Dos individuos i y j forman una *pareja* $\mathcal{P}$*-bloqueadora* para el emparejamiento μ si $j \succ_i \mu(i)$ e $i \succ_j \mu(j)$. μ es $\mathcal{P}$*-estable* si no tiene parejas $\mathcal{P}$-bloqueadoras.

El *problema de los compañeros de cuarto* consiste, pues, en determinar si para un perfil de preferencias $\mathcal{P}$ existen emparejamientos estables, y hallarlos cuando es así.

El problema de los compañeros de cuarto puede considerarse una extensión del problema del emparejamiento estable: dados H y M, puede escogerse $I = T$ (como en la definición 3.6), y se define $i \succ_i j$ siempre que $i \neq j$ sean del mismo género. Los teoremas 2.7 y 2.15 garantizan entonces que un emparejamiento μ entre H y M es estable (en el sentido de la subsección 2.1.2) si y sólo si su restricción a T es un emparejamiento estable en el sentido de los compañeros de cuarto y $\mu(i) = i$ para todo $i \notin T$.

Esto sugiere una estrategia para resolver el problema de la equidad de género en emparejamientos estables: si se posee un algoritmo para resolver el problema de los compañeros de cuarto, éste puede aplicarse al perfil de preferencias sobre T como se acaba de definir. Puesto que el algoritmo no usa información sobre el género de los participantes, debería satisfacer indiferencia de género.

Esta estrategia enfrenta, sin embargo, dos dificultades. La primera es que el problema de los compañeros de cuarto en general no tiene solución. Gale y Shapley [1] plantean las siguientes preferencias entre cuatro individuos a, b, c, d:

a	b	c	d
b	c	a	b
c	a	b	c
d	d	d	a

Entonces cualquier emparejamiento μ es bloqueado:

- Si $\mu(a) = b$, entonces b y c bloquean μ;
- si $\mu(a) = c$, entonces a y b bloquean μ;
- si $\mu(a) = d$, entonces a y c bloquean μ.

La segunda dificultad que se hace evidente es que, al satisfacer indiferencia de género, un algoritmo para el problema de los compañeros de cuarto necesariamente incumple el principio de indiferencia entre pares, por el teorema 3.5.

A pesar de estas dificultades, es provechoso estudiar el problema de los compañeros de cuarto en el contexto de la equidad de género, ya que da una idea del tipo de estrategias de solución que pueden o no dar resultado.

3.3.2. Solución al problema

Irving [7] propone un algoritmo que determina si un perfil de preferencias $\mathcal{P}$ admite un emparejamiento estable, e identifica uno si es así. El procedimiento puede esbozarse como sigue:

1. Se ejecuta primero una ronda de propuestas y rechazos al modo de Gale–Shapley, en la cual todos los participantes proponen y rechazan;

2. se reducen las listas de preferencias eliminando de manera recíproca aquellos individuos que cada participante considera inferiores a su mejor propuesta;

3. si todas las listas de preferencias constan de un solo individuo o si alguna lista de preferencias se ha quedado vacía, terminar;

4. de lo contrario, localizar una *rotación*, eliminarla y repetir desde el paso 3.

En el paso 4, una rotación es una sucesión de r parejas de individuos distintos

$$(i_0, j_0), (i_1, j_1), \ldots, (i_{r-1}, j_{r-1}),$$

tales que en la lista (reducida) de i_s, j_s ocupa el primer lugar y j_{s+1} ocupa el segundo lugar (los subíndices son módulo r), para $0 \leq s \leq r-1$. Eliminar la rotación consiste en eliminar recíprocamente a j_s de la lista de i_s, lo cual en efecto implica que cada i_s le propone a j_{s+1}.

Así pues, si en el paso 3 todas las listas constan de un solo individuo, las listas especifican un emparejamiento estable. Si en cambio la lista de un individuo i se ha quedado vacía, entonces i no puede emparejarse de manera estable y por tanto no existen emparejamientos estables para el perfil de preferencias dado.

Es claro que el algoritmo no satisface indiferencia entre pares, ya que al momento de localizar una rotación es posible que haya más de una, y el algoritmo debe seleccionar una de forma esencialmente arbitraria con respecto a la estructura de preferencias. Sin embargo, la estructura general del algoritmo puede considerarse una motivación para el algoritmo de la sección 4.3, que sí tendrá indiferencia entre pares además de indiferencia de género.

Capítulo 4

Una solución con equidad de género

En el capítulo anterior pudo comprobarse que la equidad de género en el problema del emparejamiento estable es inalcanzable en un sentido muy esencial. Una pregunta natural es si la noción de estabilidad puede generalizarse a un concepto de solución que sí admita equidad de género. En el presente capítulo se introduce tal concepto de solución y se ofrece un método de calcularlo a partir de cualquier perfil de preferencias.

4.1. Emparejamientos multivaluados

Considérese el perfil de preferencias entre dos hombres y dos mujeres de la página 23:

$$\begin{array}{cccc} \alpha & \beta & A & B \\ \hline A & B & \beta & \alpha \\ B & A & \alpha & \beta \end{array}$$

De manera intuitiva, la única "solución" que es indiferente al género consiste en declarar que ambos emparejamientos son igualmente válidos; en efecto, que tiene sentido emparejar a cada participante con ambas personas del género opuesto simultáneamente.

4.1.1. Antecedentes

En esta sección se propone una noción de emparejamiento que permite que cada persona tenga más de una pareja a la vez. Esto hace pertinente mencionar los trabajos anteriores donde se desarrolla esta posibilidad [8, 9, 12, 17, 18].

La opción de asignar múltiples parejas a un mismo individuo se remonta al origen mismo del problema del emparejamiento estable. La motivación para el artículo original de Gale y Shapley [1] involucra procesos de admisión de estudiantes a centros educativos. En su modelo, los estudiantes se identifican con los hombres; las universidades, con las mujeres; y aunque cada estudiante sólo puede asistir a una universidad, cada universidad puede por supuesto recibir a más de un estudiante a la vez. El modelo es entonces *uno-a-muchos*; en efecto, cada universidad U tiene una cota máxima q_U de estudiantes admitidos. El algoritmo se formula inicialmente para el caso *uno-a-uno* en que $q_U = 1$ para toda U, y se extiende sin dificultad alguna para q_U arbitrarios. La noción de bloqueo ahora requiere que una universidad tenga menos de q_U admitidos o que esté dispuesta a abandonar a uno de sus q_U admitidos para recibir a un aspirante que considera mejor.

El algoritmo de Gale y Shapley para el caso uno-a-uno se aplica sin problemas al caso uno-a-muchos empleando el algoritmo *estudiantes-proponen* y permitiendo que la universidad U conserve en cada iteración sus mejores q_U aspirantes. El resultado así obtenido es estable y estudiante-óptimo. Sin embargo, Roth [9] muestra que puede haber asignaciones no estables que toda universidad prefiere sobre la mejor asignación estable para las universidades, cosa que no ocurre en el modelo uno-a-uno.

La dificultad fundamental en los casos uno-a-muchos y *muchos-a-muchos* es que se hace necesario describir las preferencias de los participantes no sólo con respecto a individuos, sino con respecto a conjuntos de individuos [9]. Esto obliga a especificar los órdenes de preferencias sobre conjuntos, y a definir nociones de estabilidad que incorporen estas preferencias [17]:

- Un emparejamiento es *estable por parejas* si no hay un hombre y una mujer que, no estando juntos, preferirían estarlo (posiblemente abandonando o conservando otras de sus respectivas parejas);

- un emparejamiento está en el *núcleo* si no existe una coalición de hombres y mujeres que, emparejándose sólo entre ellos, obtengan un resultado mejor que lo que el emparejamiento les asigna;

- un emparejamiento es *estable por conjuntos* si no existe una coalición de hombres y mujeres que, agregando vínculos entre ellos (posiblemente abandonando o conservando otras de sus respectivas parejas), puedan obtener un resultado mejor que lo que el emparejamiento les asigna.

Sotomayor [17] muestra que la estabilidad por conjuntos es estrictamente más fuerte que la estabilidad por parejas y el núcleo; muestra también que para una estructura de preferencias el núcleo puede no contener emparejamientos estables por parejas, y por tanto los emparejamientos estables por conjuntos pueden no existir.

Cabe notar que este planteamiento con preferencias sobre conjuntos es efectivamente una formulación variante del problema, fruto del hecho de que los emparejamientos muchos-a-muchos son en este caso la meta, y por tanto la noción de estabilidad debe genuinamente reflejar las opciones estratégicas que enfrentan los jugadores. Los conceptos que se presentan en esta sección, en cambio, cobran sentido sólo como un paso intermedio en la obtención de uno o más emparejamientos estables en el sentido clásico, y por tanto deben prescindir de cualquier estructura adicional a los órdenes de preferencias sobre individuos.

4.1.2. Motivación

Supóngase que, para cada perfil de preferencias $\mathcal{P}$ entre conjuntos H y M, quiere obtenerse una subcolección de emparejamientos estables

$$\mathcal{E}(H, M, \mathcal{P}) \subseteq \mathrm{Est}(\mathcal{P}).$$

Supóngase además que estas colecciones quieren obtenerse con un procedimiento que satisfaga criterios análogos a la indiferencia de género e indiferencia entre pares de la sección 3.1:

$$\mathcal{E}(M, H, \mathcal{P}) = \mathcal{E}(H, M, \mathcal{P}),$$
$$\mathcal{E}(\theta(H), \theta(M), \theta\mathcal{P}) = \{\theta\mu \mid \mu \in \mathcal{E}(H, M, \mathcal{P})\}.$$

En algunos casos será inevitable, para cumplir estos criterios, que $\mathcal{E}(\mathcal{P}) = \mathrm{Est}(\mathcal{P})$, por ejemplo para el perfil de preferencias de la página 23. Sin embargo, en otros casos habrá una forma más eficiente de asegurar los dos criterios;

considérese esta lista de preferencias $\mathcal{P}$:

α	β	γ	A	B	C
A	B	C	β	γ	α
B	C	A	γ	α	β
C	A	B	α	β	γ

Puede determinarse que

$$\mathrm{Est}(\mathcal{P}) = \{\mu_H, \mu_M, \mu_2\},$$

donde μ_H y μ_M le dan respectivamente a hombres y mujeres sus parejas preferidas, y μ_2 le da a todas las personas su segunda pareja preferida. Entonces cualquier permutación que fije $H \cup M$ y fije la estructura de preferencias fija también el emparejamiento μ_2, de forma que elegir

$$\mathcal{E}(H, M, \mathcal{P}) = \{\mu_2\}$$

es compatible con ambos criterios. Esta elección tiene la propiedad adicional de que ningún individuo queda emparejado jamás con su peor pareja factible, sino a lo sumo con su segunda peor.

Surge, pues, una pregunta natural: ¿cuánto más pequeño puede esperarse que sea el conjunto $\mathcal{E}(H, M, \mathcal{P})$? Mejor dicho, ¿cuántos emparejamientos pueden excluirse de $\mathcal{E}(H, M, \mathcal{P})$ sin romper los criterios de justicia? En esta sección se introducirá una extensión natural de la noción de estabilidad que justifica la idea de que, para cualquier perfil de preferencias $\mathcal{P}$, basta considerar

$$|\mathcal{E}(H, M, \mathcal{P})| \leq 2.$$

4.1.3. Definiciones

Definición 4.1. Un *emparejamiento multivaluado* entre H y M es una función

$$\lambda : H \cup M \to \wp(H \cup M)$$

que satisface, para todo $h \in H$, $m \in M$:

- $\lambda(h), \lambda(m) \neq \varnothing$;
- $\lambda(h) \subseteq M \cup \{h\}$, $\lambda(m) \subseteq H \cup \{m\}$;

- *simetría*: $h \in \lambda(m)$ si y sólo si $m \in \lambda(h)$;
- *monogamia mutua*: $\lambda(h) = \{m\}$ si y sólo si $\lambda(m) = \{h\}$.

La condición de monogamia mutua no es imprescindible si lo que se quiere es emparejar a cada participante con más de una persona a la vez, pero se justifica en un objetivo pragmático y en una interpretación del modelo:

- El objetivo es obtener emparejamientos uno-a-uno a partir de un emparejamiento multivaluado, lo cual puede no ser posible si no se satisface monogamia mutua. Si por ejemplo $\lambda(h_1) = \{m\}$ pero $\lambda(m) = \{h_1, h_2\}$, entonces emparejar a m con h_2 haría imposible definir una pareja para h_1. En el peor de los casos, debe escogerse $\lambda(h_1) = \{m, h_1\}$, haciendo explícita la opción de dejar soltero a h_1.

- Bajo los supuestos de que cada individuo distribuye su "tiempo" entre distintas personas, y de que todas las personas disponen de la misma cantidad de tiempo para distribuir, es válido interpretar $\lambda(i)$ como el conjunto de personas con las cuales i pasa un tiempo estrictamente positivo. Ahora bien, si $\lambda(h_1) = \{m\}$ pero $\lambda(m) = \{h_1, h_2\}$, esto querría decir que h_1 pasa el 100 % de su tiempo con m, pero ella no pasa el 100 % de su tiempo con h_1, lo cual no tiene sentido.

Antes de introducir la noción de estabilidad para emparejamientos multivaluados, es propio hacer una distinción. Formalmente, un emparejamiento $\mu : H \cup M \to H \cup M$ no es lo mismo que un emparejamiento multivaluado λ, incluso si resulta ser que $|\lambda(i)| = 1$ para todo i, pues las imágenes de λ son conjuntos de individuos, mientras que las imágenes de μ son individuos en sí. De manera preliminar, los emparejamientos $\mu : H \cup M \to H \cup M$ se denominarán *emparejamientos univaluados* para distinguirlos del siguiente concepto.

Definición 4.2. Un emparejamiento multivaluado $\lambda : H \cup M \to \wp(H \cup M)$ es *monógamo* si $|\lambda(i)| = 1$ para todo $i \in H \cup M$.

Así pues, un emparejamiento univaluado envía personas en personas, mientras que un emparejamiento multivaluado monógamo envía personas en conjuntos unitarios de personas. Obsérvese que un emparejamiento univaluado μ siempre define un emparejamiento monógamo λ por $\lambda(i) = \{\mu(i)\}$ para todo i, y recíprocamente, la misma expresión define un emparejamiento univaluado μ a partir de un emparejamiento multivaluado λ. Esta distinción, sin embargo, resulta ser meramente formal: con respecto a la estabilidad, los emparejamientos monógamos extienden precisamente a los emparejamientos univaluados, como se verá.

4.1.4. Estabilidad en emparejamientos multivaluados

Las nociones de bloqueo y de estabilidad se extienden de manera natural a los emparejamientos multivaluados.

Definición 4.3. Sea $\mathcal{P}$ un perfil de preferencias entre H y M. Los individuos $i, j \in H \cup M$ *bloquean* un emparejamiento multivaluado λ si existen $i_\lambda \in \lambda(i)$ y $j_\lambda \in \lambda(j)$ tales que

$$j \succ_i i_\lambda,$$
$$i \succ_j j_\lambda.$$

Si λ no es bloqueado, entonces es *estable*.

Como en la estabilidad ordinaria, si $i = j$ estas condiciones se reducen a que existe $i_\lambda \in \lambda(i)$ tal que i prefiere estar soltero que estar con i_λ.

Esta definición extiende la noción de estabilidad para emparejamientos univaluados, en el sentido de que dos individuos i y j bloquean un emparejamiento μ si y sólo si bloquean el emparejamiento monógamo λ obtenido a partir de μ como se discutió antes, pues la condición $i_\lambda \in \lambda(i)$ equivale a $i_\lambda = \mu(i)$ dado que

$$\lambda(i) = \{\mu(i)\}.$$

En consecuencia, μ es estable (en el sentido univaluado) si y sólo si λ es estable (en el sentido multivaluado). Así pues, en lo sucesivo no se distinguirá entre emparejamientos univaluados y emparejamientos monógamos.

Ejemplo. Considérese la lista de preferencias $\mathcal{P}$ entre dos hombres α, β y dos mujeres A, B:

α	β	A	B
A	B	β	α
B	A	α	β

En el ejemplo de la página 23 se observó que $\mathcal{P}$ tiene dos emparejamientos estables μ_H y μ_M. Desde luego ambos son estables como emparejamientos univaluados. También hay un emparejamiento multivaluado estable λ dado por

$$\lambda(\alpha) = \lambda(\beta) = M, \qquad \lambda(A) = \lambda(B) = H.$$

Ahora bien, λ puede considerarse la unión de μ_H y μ_M en el sentido de que

$$\lambda(i) = \{\mu_H(i), \mu_M(i)\} \text{ para todo } i \in H \cup M.$$

En este caso, la unión de los dos emparejamientos univaluados estables da lugar a un emparejamiento multivaluado estable. ¿Será éste siempre el caso? A continuación se verá que no.

Definición 4.4. Sean λ_1, λ_2 emparejamientos multivaluados entre H y M. Se dice que λ_1 es un *subemparejamiento* de λ_2, y se escribe $\lambda_1 \subseteq \lambda_2$, si $\lambda_1(i) \subseteq \lambda_2(i)$ para todo $i \in H \cup M$.

Teorema 4.5. Sean λ un emparejamiento multivaluado estable entre H y M, y sea κ un subemparejamiento de λ. Entonces κ es estable.

Demostración. Supóngase a modo de contradicción que i y j bloquean κ. Entonces existen $i_\kappa \in \kappa(i)$ y $j_\kappa \in \kappa(j)$ tales que

$$j \succ_i i_\kappa, \qquad i \succ_j j_\kappa.$$

Dado que $\kappa \subseteq \lambda$, $i_\kappa \in \lambda(i)$ y $j_\kappa \in \lambda(j)$. Por tanto, i y j también bloquean λ, contradiciendo que λ es estable. Se concluye que κ no tiene bloqueos. $\square$

Ejemplo. Considérese la lista de preferencias de la página 36:

α	β	γ	A	B	C
A	B	C	β	γ	α
B	C	A	γ	α	β
C	A	B	α	β	γ

Entonces $\lambda = \mu_H \cup \mu_M$ no es estable, pues

$$B \succ_\alpha C \in \lambda(\alpha),$$
$$\alpha \succ_B \beta \in \lambda(B),$$

es decir, α y B bloquean λ. Por el teorema 4.5, la unión de todos los emparejamientos univaluados estables tampoco es un emparejamiento multivaluado estable.

Se obtiene, pues, que la estabilidad multivaluada extiende la estabilidad univaluada, pero no de una manera enteramente trivial. A continuación se desarrollan las propiedades de los emparejamientos multivaluados y la relación con sus subemparejamientos univaluados.

4.1.5. Emparejamientos univaluados a partir de emparejamientos multivaluados

Teorema 4.6. Dado un emparejamiento multivaluado estable λ entre H y M, defínase una función $f_\lambda : H \cup M \to H \cup M$ por

$$f_\lambda(i) = \max_{\mathcal{P}i} \lambda(i), \text{ para cada } i \in H \cup M.$$

Entonces f_λ es una biyección.

Demostración. Puesto que $H \cup M$ es finito, basta demostrar que f_λ es inyectiva. A modo de contradicción supóngase que no; sean $i, j, k \in H \cup M$ tales que

$$f_\lambda(i) = f_\lambda(j) = k, \qquad\qquad i \neq j.$$

Sin pérdida de generalidad, $i \succ_k j$. Puesto que $\{i, j\} \subset \lambda(k)$, por monogamia mutua existe $\ell \in \lambda(i)$ distinto de k. Entonces

$$\begin{aligned} i \succ_k j &\in \lambda(k), \\ k \succ_i \ell &\in \lambda(i), \end{aligned}$$

esto último ya que $k = \max_{\mathcal{P}i} \lambda(i)$. Por tanto, i y k bloquean λ, contradiciendo su estabilidad. Esto prueba que f_λ es inyectiva. □

Teorema 4.7. Defínase $\mu : H \cup M \to H \cup M$ por

$$\mu(i) = \begin{cases} f_\lambda(i) & \text{si } i \in H, \\ f_\lambda^{-1}(i) & \text{si } i \in M. \end{cases}$$

Entonces μ es un emparejamiento, $\mu \subseteq \lambda$ y μ es estable.

Demostración. Obsérvese primero que

$$\begin{aligned} \mu(h) = f_\lambda(h) \in \lambda(h) \subset M \cup \{h\} &\qquad \text{para } h \in H, \\ \mu(m) = f_\lambda^{-1}(m) \in \lambda(m) \subset H \cup \{m\} &\qquad \text{para } m \in M. \end{aligned}$$

Entonces, para $i \in H \cup M$,

$$\mu(\mu(i)) = \begin{cases} f_\lambda^{-1}(f_\lambda(i)) & \text{si } i \in H \text{ y } \mu(i) \in M, \\ i & \text{si } i \in H \text{ y } \mu(i) = i, \\ f_\lambda(f_\lambda^{-1}(i)) & \text{si } i \in M \text{ y } \mu(i) \in H, \\ i & \text{si } i \in M \text{ y } \mu(i) = i. \end{cases}$$

Por tanto, $\mu(\mu(i)) = i$ en todos los casos. Esto demuestra que μ es un emparejamiento. Desde luego, $\mu \subseteq \lambda$ como se acaba de observar. La estabilidad de μ se sigue entonces del teorema 4.5. □

Teorema 4.8. Sea λ un emparejamiento multivaluado estable entre H y M, y sea $h \in H$. Entonces $|\lambda(h)| \leq 2$.

Demostración. Supóngase a modo de contradicción que existen m_1, m_2, m_3 distintas en $\lambda(h)$, y sin pérdida de generalidad

$$m_1 \succ_h m_2 \succ_h m_3.$$

Sea μ el emparejamiento del teorema 4.7, de forma que

$$\mu(k) = \max_{\mathcal{P}k} \lambda(k) \text{ para todo hombre } k \in H.$$

En consecuencia, $\mu(h) \neq m_2$. Sea $k = \mu(m_2)$. Por monogamia mutua, puesto que $\lambda(m_2) \neq \{k\}$, existe $j \in \lambda(k)$ distinta de m_2, y por tanto

$$m_2 \succ_k j \in \lambda(k).$$

Dado que λ es estable, debe tenerse

$$h \succ_{m_2} k \in \lambda(m_2),$$

pues de lo contrario m y k bloquearían λ. Por otro lado,

$$m_2 \succ_h m_3 \in \lambda(h),$$

lo que significa que h y m bloquean λ. Esto es una contradicción. □

Si bien un emparejamiento multivaluado estable λ puede contener más de dos emparejamientos univaluados, los resultados anteriores garantizan que λ puede expresarse como unión de solamente dos emparejamientos univaluados estables λ_H y λ_M, a saber,

$$\lambda_H = \max_{\geq_H} \{\mu \mid \mu \subseteq \lambda\},$$
$$\lambda_M = \max_{\geq_M} \{\mu \mid \mu \subseteq \lambda\},$$

donde $\leq_H$ y $\leq_M$ son los órdenes parciales del teorema 2.17.

En particular, λ queda especificado tan pronto como se sabe cuáles son las dos (o menos) parejas de cada persona. Esto sugiere que una forma de localizar un emparejamiento multivaluado estable es partir del perfil de preferencias original y, eliminando parejas bloqueadas, reducir el perfil hasta que solamente queden dos personas o menos en la lista de cada participante.

En principio, un algoritmo que desarrolle esta idea conservando estabilidad, indiferencia de género e indiferencia entre pares podrá considerarse una solución al problema. En las siguientes dos subsecciones se propone un procedimiento de este tipo.

4.2. El algoritmo de Gale–Shapley bilateral

Un paso intermedio en el desarrollo del procedimiento consiste en reducir el perfil de preferencias de tal manera que el conjunto de emparejamientos estables permanezca invariante. En esta sección se muestra cómo conseguirlo.

4.2.1. Operaciones con listas de preferencias

Como se vio en la subsección 2.1.3, $\mathrm{Est}(\mathcal{P})$ está bien definido para cada lista de preferencias $\mathcal{P}$. Se define ahora otro invariante con respecto a las listas de preferencias:

Definición 4.9. Sea $\mathcal{P}$ un perfil de preferencias entre H y M. El *conjunto de parejas $\mathcal{P}$-admisibles* de $i \in H \cup M$ se denota por

$$\mathcal{P}[i] = \{j \in H \cup M \mid j \succ_i i \text{ en } \mathcal{P}\}.$$

Por la definición de la relación de equivalencia, $\mathcal{P}[i]$ está bien definido para cada lista de preferencias $\mathcal{P}$ y cada i. El teorema 2.7 afirma, pues, que $\mathrm{Est}(\mathcal{P})$ queda determinado únicamente por las preferencias de cada individuo i restringidas a $\mathcal{P}[i]$.

Una operación natural con una lista de preferencias consiste en borrar un elemento de alguna columna. Considérese la siguiente lista de preferencias $\mathcal{P}$:

α	β	γ	δ	A	B	C	D
A	B	C	D	β	γ	δ	α
C	C	A	B	γ	δ	β	δ
D	A	B	C	α	β	α	
						γ	

Nótese que todos excepto C tienen una pareja descalificada; D tiene dos. Supóngase que quiere eliminarse a δ de la lista de preferencias de C. No es difícil construir un perfil de preferencias $\mathcal{Q}$ cuya tabla refleja ese cambio:

- Para $i \neq C$, el orden de preferencias $\mathcal{Q}i$ es el mismo $\mathcal{P}i$;
- el orden de preferencias $\mathcal{Q}C$ coincide con $\mathcal{P}C$ en $\{\alpha, \beta, \gamma, C\}$ y por demás pone $j \succ_C \delta$ para todo $j \in \{\alpha, \beta, \gamma, C\}$.

La lista de preferencias correspondiente a $\mathcal{Q}$ queda entonces como sigue,

α	β	γ	δ	A	B	C	D
A	B	C	D	β	γ	β	α
C	C	A	B	γ	δ	α	δ
D	A	B	C	α	β	γ	

y se dice que, al pasar de $\mathcal{P}$ a $\mathcal{Q}$, C ha "eliminado" a B, o que ha "descalificado" a B, de su lista de preferencias.

Desde luego, este tipo de descalificaciones no puede hacerse de manera arbitraria, so riesgo de eliminar algún emparejamiento estable o introducir otro que no lo era. Los siguientes resultados muestran dos operaciones con listas de preferencias que preservan emparejamientos estables.

Teorema 4.10. Sea $\mathcal{P}$ una lista de preferencias. Supóngase que i es pareja descalificada de j. Si $\mathcal{Q}$ es la lista de preferencias obtenida a partir de $\mathcal{P}$ haciendo que i también descalifique a j, entonces $\mathrm{Est}(\mathcal{Q}) = \mathrm{Est}(\mathcal{P})$.

Demostración. Supóngase primero que $\mu \in \mathrm{Est}(\mathcal{P})$. Claramente se satisfacen las hipótesis a) y b) del teorema 2.6. Para concluir que $\mu \in \mathrm{Est}(\mathcal{Q})$ basta comprobar la hipótesis c). Si $\mu(i) \succ_i k$ en $\mathcal{P}$, entonces también $\mu(i) \succ_i k$ en $\mathcal{Q}$ a menos que $\mu(i) = j$, lo cual es imposible ya que μ es $\mathcal{P}$-estable pero i no es pareja $\mathcal{P}$-admisible de j. Esto implica la hipótesis c) del teorema, de lo cual se sigue que $\mathrm{Est}(\mathcal{P}) \subseteq \mathrm{Est}(\mathcal{Q})$.

Supóngase ahora que $\mu \notin \mathrm{Est}(\mathcal{P})$. Supóngase que $k, \ell \in H \cup M$ bloquean μ con respecto a $\mathcal{P}$,

$$k \succ_\ell \mu(\ell) \text{ en } \mathcal{P}, \qquad \ell \succ_k \mu(k) \text{ en } \mathcal{P}.$$

Ahora bien, $\mathcal{P}k = \mathcal{Q}k$ y $\mathcal{P}\ell = \mathcal{Q}\ell$, y por tanto k y ℓ $\mathcal{Q}$-bloquean μ, a menos que $\{k, l\} = \{i, j\}$. Pero eso es imposible, ya que i es pareja $\mathcal{P}$-descalificada de j, y por tanto

$$\mu(j) \succ_j j \succ_j i \text{ en } \mathcal{Q}.$$

Se concluye que $\mu \notin \mathrm{Est}(\mathcal{Q})$, lo cual completa la prueba. □

Ya que estas eliminaciones pueden repetirse para cada pareja descalificada, este resultado demuestra que siempre pueden eliminarse "recíprocamente" de la lista aquellas preferencias que no sean "correspondidas", de forma que queden únicamente preferencias mutuamente admisibles. El siguiente resultado muestra hasta dónde se pueden truncar las listas de preferencias sin alterar el conjunto de emparejamientos estables.

Teorema 4.11. Sea $\mathcal{P}$ una lista de preferencias entre H y M. Sean $i \in H \cup M$, $j = \min \mathcal{P}[i]$ la peor pareja $\mathcal{P}$-admisible de i, y supóngase que i tiene una pareja $\mathcal{P}$-factible $\ell \neq j$. Obténgase $\mathcal{Q}$ haciendo que i descalifique a j. Entonces

$$\operatorname{Est}(\mathcal{Q}) = \{\mu \in \operatorname{Est}(\mathcal{P}) \mid \mu(i) \neq j\}.$$

Demostración. Supóngase primero que $\mu \in \operatorname{Est}(\mathcal{P})$ y $\mu(i) \neq j$. Para usar el teorema 2.6 ya se tienen las hipótesis a) y b). Basta comprobar la hipótesis c). Si $\mu(i) \succ_i k$ en $\mathcal{P}$, entonces $\mu(i) \succ_i k$ en $\mathcal{Q}$, ya que al pasar de $\mathcal{P}$ a $\mathcal{Q}$ el único individuo que decrece en la lista es j, y $\mu(i) \neq j$ por hipótesis. Por tanto, la hipótesis c) del teorema también se cumple. Esto prueba que μ es $\mathcal{Q}$-estable, y en consecuencia

$$\{\mu \in \operatorname{Est}(\mathcal{P}) \mid \mu(i) \neq j\} \subseteq \operatorname{Est}(\mathcal{Q}).$$

Ahora bien, i tiene una pareja $\mathcal{P}$-factible $\ell \neq j$. Si ν es un emparejamiento $\mathcal{P}$-estable con $\nu(i) = \ell$, entonces $\nu \in \operatorname{Est}(\mathcal{Q})$ como se acaba de demostrar, y el teorema 2.15 prueba que $\mu(i) \neq i$ para todo $\mu \in \operatorname{Est}(\mathcal{Q})$.

Para finalizar, supóngase que $\mu \in \operatorname{Est}(\mathcal{Q})$. Se usará nuevamente el teorema 2.6, esta vez intercambiando $\mathcal{P}$ y $\mathcal{Q}$. Las hipótesis a) y b) ya se cumplen. Para verificar la hipótesis c), nótese que $\mu(i) \neq i$ como se acaba de observar, y $\mu(i) \neq j$ ya que j es pareja $\mathcal{Q}$-descalificada de i. Por tanto,

$$\mu(i) \succ_i i \succ_i j \text{ en } \mathcal{Q}.$$

En particular, $\mu(i)$ es $\mathcal{P}$-admisible para i. Como j es la peor pareja $\mathcal{P}$-admisible de i,

$$\mu(i) \succ_i j \text{ en } \mathcal{P}.$$

Se concluye que si $\mu(i) \succ_i k$ en $\mathcal{Q}$, entonces también $\mu(i) \succ_i k$ en $\mathcal{P}$. Esto es cierto cuando $k \neq j$ porque las preferencias coinciden, y también cuando $k = j$ por lo que se acaba de observar. Por tanto, la hipótesis c) también se cumple, y esto demuestra que

$$\operatorname{Est}(\mathcal{Q}) \subseteq \{\mu \in \operatorname{Est}(\mathcal{P}) \mid \mu(i) \neq j\}. \qquad \square$$

De acuerdo con este resultado, un individuo puede descalificar a cualquier cantidad de parejas potenciales sin quedarse soltero, siempre y cuando no elimine a su mejor pareja factible. Esto será crucial en la sección 4.3, pero de manera inmediata es más útil la siguiente consecuencia.

Teorema 4.12. Si en el teorema 4.11 j no es una pareja factible de i, entonces $\mathrm{Est}(\mathcal{Q}) = \mathrm{Est}(\mathcal{P})$.

Demostración. En tal caso, $\{\mu \in \mathrm{Est}(\mathcal{P}) \mid \mu(i) \neq j\} = \mathrm{Est}(\mathcal{P})$. □

Este resultado demuestra que un individuo i puede descalificar a aquellas parejas que considera peores que su peor pareja factible, y eso no modifica el conjunto de emparejamientos estables.

4.2.2. Formulación del algoritmo

Los teoremas 4.10 y 4.12 muestran hasta dónde se puede reducir $\mathcal{P}$ con la tranquilidad de preservar $\mathrm{Est}(\mathcal{P})$. Esto sugiere el siguiente procedimiento para reducir $\mathcal{P}$ a una lista $\mathcal{P}^*$ con los mismos emparejamientos estables:

1. localizar la peor pareja $\mathcal{P}$-factible $\mathrm{PPF}(i, \mathcal{P})$ para cada individuo i;
2. truncar la lista de preferencias de cada i por debajo de $\mathrm{PPF}(i, \mathcal{P})$;
3. eliminar todas las preferencias que no sean mutuamente admisibles.

El teorema 4.12 garantiza que el paso 2 preserva $\mathrm{Est}(\mathcal{P})$, y el teorema 4.10 implica que, tras el paso 3, $\mathrm{Est}(\mathcal{P}^*) = \mathrm{Est}(\mathcal{P})$.

En el paso 1, localizar la peor pareja factible de i puede hacerse por cualquier método, y en particular usando el algoritmo de Gale–Shapley tal que el género que propone es el género opuesto al de i. Entonces el teorema 2.17 garantiza que la pareja de i es su peor pareja factible.

Teorema 4.13. Para todo $i, j \in H \cup M$ distintos, $j \in \mathcal{P}^*[i]$ si y sólo si (en el perfil $\mathcal{P}$)

$$j \succeq_i \mathrm{PPF}(i, \mathcal{P}) \qquad \text{y} \qquad i \succeq_j \mathrm{PPF}(j, \mathcal{P}).$$

Además, si i no es soltero en $\mathcal{P}$, entonces $\mathrm{MPF}(i, \mathcal{P}) = \text{máx}\, \mathcal{P}^*[i]$ (relativo al orden $\mathcal{P}i$).

Demostración. Si $j \succeq_i \mathrm{PPF}(i, \mathcal{P})$ e $i \succeq_j \mathrm{PPF}(j, \mathcal{P})$, entonces ninguno de ellos elimina al otro en el paso 2, y en consecuencia tampoco en el paso 3. Por tanto, $j \in \mathcal{P}^*[i]$.

En cambio, si $\mathrm{PPF}(i, \mathcal{P}) \succ_i j$, entonces i elimina a j en el paso 2; y si $\mathrm{PPF}(j, \mathcal{P}) \succ_j i$, entonces j elimina a i en el paso 2 y por tanto i elimina a j en el paso 3. En ambos casos, $j \notin \mathcal{P}^*[i]$. Esto demuestra la primera parte del enunciado.

Supóngase entonces que i no es soltero en $\mathcal{P}$, y sea $k = \mathrm{MPF}(i, \mathcal{P})$. El teorema 2.17 garantiza que $i = \mathrm{PPF}(k, \mathcal{P})$, y por definición $k \succeq_i \mathrm{PPF}(i, \mathcal{P})$. Como se acaba de demostrar, $k \in \mathcal{P}^*[i]$.

Para demostrar que $k = \max \mathcal{P}^*[i]$, debe verse que si $j \succ_i k$ entonces $j \notin \mathcal{P}^*[i]$. Dado un tal j, sea $\mu \in \mathrm{Est}(\mathcal{P})$. Entonces $j \succ_i \mu(i)$, pues k es la mejor pareja factible de i. Por estabilidad, $\mu(j) \succ_j i$. Como esto es válido para cualquier $\mu \in \mathrm{Est}(\mathcal{P})$,

$$\mathrm{PPF}(j, \mathcal{P}) \succ_j i,$$

y la primera parte del enunciado implica que $j \notin \mathcal{P}^*[i]$, como se quería. □

En principio, nada impide aplicar el algoritmo de Gale–Shapley simultáneamente a ambos géneros, de lo cual resulta el nombre de *algoritmo de Gale–Shapley bilateral*. Sin embargo, este algoritmo no produce un emparejamiento μ, sino una lista de preferencias reducida.

Tal y como el algoritmo base de Gale–Shapley, este algoritmo procede por etapas. En cada etapa:

1. cada participante que no tenga una lista vacía le hace una propuesta a la primera persona de su lista;

2. cada participante que reciba al menos una propuesta descalifica a todas las personas que prefiere menos que a su mejor propuesta;

3. cada participante descalifica a todas las personas que lo han descalificado.

El algoritmo termina tan pronto como en una etapa no ocurren más descalificaciones, lo cual siempre se alcanza ya que las listas de preferencias son finitas. El resultado es una lista de preferencias $\mathcal{Q}$ en la cual cada individuo ha descalificado a algún subconjunto (posiblemente vacío) de sus parejas admisibles.

Teorema 4.14. $\mathcal{Q} = \mathcal{P}^*$.

Demostración. En el algoritmo de Gale–Shapley bilateral, un individuo i descalifica a un individuo $j \neq i$ si y sólo si ocurre uno de los siguientes casos:

a) i ha recibido una mejor propuesta de otro individuo k, de forma que

$$\mathrm{PPF}(i, \mathcal{P}) \succeq_i k \succ_i j;$$

b) i es pareja descalificada de j en $\mathcal{P}$, de forma que

$$\mathrm{PPF}(j, \mathcal{P}) \succeq_j j \succ_j i;$$

c) j descalifica a i en el desarrollo del algoritmo, lo cual equivale al caso a) con los papeles de i y j intercambiados.

En otras palabras, $j \notin \mathcal{Q}[i]$ si y sólo si $\mathrm{PPF}(i, \mathcal{P}) \succ_i j$ o $\mathrm{PPF}(j, \mathcal{P}) \succ_j i$. El teorema 4.13 da entonces la igualdad buscada. □

Teorema 4.15. Defínase $\lambda : H \cup M \to \wp(H \cup M)$ por

$$\lambda(i) = \begin{cases} \mathcal{P}^*[i] & \text{si } \mathcal{P}^*[i] \neq \varnothing, \\ \{i\} & \text{si } \mathcal{P}^*[i] = \varnothing. \end{cases}$$

Entonces λ es un emparejamiento multivaluado $\mathcal{P}$-estable si y sólo si $|\mathcal{P}^*[i]| \leq 2$ para todo i.

Demostración. Por la definición, $|\mathcal{P}^*[i]| \leq |\lambda(i)|$ para todo i, y la desigualdad es estricta sólo cuando $\mathcal{P}^*[i] = \varnothing$.

Se verifica primero que λ es un emparejamiento multivaluado. En efecto,

- $\lambda(i) \neq \varnothing$ para cada i por construcción;
- $\lambda(h) \subseteq \mathcal{P}[h] \subseteq M \cup \{h\}$ para cada $h \in H$, y $\lambda(m) \subseteq \mathcal{P}[m] \subseteq H \cup \{m\}$ para cada $m \in M$;
- λ es simétrico, ya que en el procedimiento todas las parejas que no son mutuamente admisibles son eliminadas de la lista;
- λ satisface monogamia mutua, pues si $\lambda(i) = \{j\}$ con $j \neq i$, entonces j es la única pareja factible de i (por el teorema 4.13), y por tanto $\lambda(j) = \{i\}$.

Así pues, si λ es estable entonces el teorema 4.8 implica que, para todo i,

$$|\mathcal{P}^*[i]| \leq |\lambda(i)| \leq 2.$$

Recíprocamente, si λ no es estable, entonces existen $i, j, i_\lambda, j_\lambda$ tales que (en el perfil $\mathcal{P}$)

$$j \succ_i i_\lambda \in \lambda(i), \qquad i \succ_j j_\lambda \in \lambda(j).$$

Esto implica lo siguiente:

- $j \succ_i \lambda(i) \succeq_i \mathrm{PPF}(i, \mathcal{P})$ y $i \succ_j \lambda(j) \succeq_j \mathrm{PPF}(j, \mathcal{P})$. Por el teorema 4.13,

$$j \in \lambda(i).$$

- Sea $k = \mathrm{MPF}(i, \mathcal{P}) \in \lambda(i)$ y supóngase a modo de contradicción que $j = k$. Por el teorema 2.19,

$$\mathrm{PPF}(j, \mathcal{P}) = i \succ_j j_\lambda.$$

  Entonces j_λ es descalificada por j en el desarrollo del algoritmo, contradiciendo que $j_\lambda \in \lambda(j)$. Se sigue que $j \neq k$.

- Sea $\ell = \mathrm{PPF}(i, \mathcal{P}) \in \lambda(i)$. Entonces $\ell \neq j$, pues $j \succ_i i_\lambda \succeq_i \ell$.

Resulta entonces que

$$k \succ_i j \succ_i \ell, \qquad k, j, \ell \in \mathcal{P}^*[i],$$

y por tanto

$$|\mathcal{P}^*[i]| = |\lambda(i)| \geq 3. \qquad \square$$

4.3. Un algoritmo multivaluado estable

Los resultados de la sección anterior ofrecen un medio para obtener un emparejamiento multivaluado estable: después de aplicar el algoritmo de Gale–Shapley bilateral, si alguna persona i todavía tiene más de dos personas en su lista, el teorema 4.11 garantiza que i puede descalificar a su peor pareja factible y el conjunto de emparejamientos estables se reduce. Tan pronto como todos los participantes tienen dos o menos personas en su lista, el teorema 4.15 garantiza que se ha obtenido un emparejamiento multivaluado estable.

4.3.1. Truncar parejas estables

Es posible, desde luego, que haya más de una persona con más de tres parejas en lista. La tarea siguiente es determinar cuál de ellas, o cuáles, deben eliminar a su peor pareja factible. Si bien cualquiera de las eliminaciones conducirá en última instancia a un emparejamiento estable, esta elección no puede ser enteramente arbitraria, pues eso violaría el criterio de indiferencia entre pares.

Ejemplo. La solución ideal consistiría en permitir que toda persona i con $|\mathcal{P}^*[i]| \geq 3$ elimine a su peor pareja factible. El siguiente ejemplo muestra que esa elección puede no conducir a un emparejamiento estable. Considérese la siguiente lista de preferencias $\mathcal{P}$ entre $H = \{\alpha, \beta, \gamma, \delta\}$ y $M = \{A, B, C, D\}$:

α	β	γ	δ	A	B	C	D
A	B	C	D	β	γ	δ	α
C	C	A	B	γ	δ	β	δ
D	A	B	C	α	β	α	
						γ	

Esta lista ya está reducida hasta la peor pareja factible de cada individuo, de forma que $\mathcal{P}^* = \mathcal{P}$. Así pues, $|\mathcal{P}^*[i]| \geq 3$ para todo participante i, excepto para δ; si se permite que todos menos δ eliminen a su peor pareja factible, se obtiene la siguiente lista:

α	β	γ	δ	A	B	C	D
A	B	C	D	β	γ	δ	α
C	C	A	B	γ	δ	β	δ
						α	

Sin embargo, ejecutar el algoritmo de Gale–Shapley *hombres-proponen* con esta lista produce el emparejamiento μ que asigna

$$\mu(\alpha) = \alpha, \qquad \mu(\beta) = C, \qquad \mu(\gamma) = A, \qquad \mu(\delta) = D, \qquad \mu(A) = A,$$

que no es $\mathcal{P}$-estable, ya que α y A prefieren estar juntos que estar solteros. Permitir que se hagan todas las eliminaciones simultáneamente ha hecho que algunas personas queden solteras cuando no debían estarlo.

En el siguiente resultado se da una forma de seleccionar las eliminaciones de manera que no se deje indebidamente soltero a ningún participante.

Teorema 4.16. Sea $\mathcal{Q}$ la lista de preferencias obtenida a partir de $\mathcal{P}$ haciendo

$$\mathcal{Q}[i] = \mathcal{P}^*[i] \setminus \{\mathrm{PPF}(i, \mathcal{P})\}.$$

Sea $T = \{i \in H \cup M \mid i \text{ no es soltero de } \mathcal{Q}\}$. Entonces existe un emparejamiento $\mu \in \mathrm{Est}(\mathcal{P})$ tal que (nótese que la preferencia es estricta)

$$\mu(i) \succ_i \mathrm{PPF}(i, \mathcal{P}) \text{ en } \mathcal{P}, \qquad \text{para todo } i \in T.$$

Demostración. Para empezar, obsérvese que si $i \in T$, entonces $|\mathcal{P}^*[i]| \geq 2$, ya que rechazar a su peor pareja factible no ha hecho que i se quede soltero. En particular, si $i \in T$,

$$\mathrm{MPF}(i, \mathcal{P}) \succ_i \mathrm{PPF}(i, \mathcal{P}) \text{ en } \mathcal{P}.$$

Sea $\mathcal{R}$ la lista de preferencias obtenida a partir de $\mathcal{P}^*$ permitiendo que cada $m \in M \cap T$ elimine a su peor pareja $\mathcal{P}$-factible. Éste se puede obtener a partir de $\mathcal{P}$ por eliminaciones sucesivas, de forma que, por el teorema 4.11,

$$\mathrm{Est}(\mathcal{R}) = \{\mu \in \mathrm{Est}(\mathcal{P}) \mid \mu(m) \neq \mathrm{PPF}(m, \mathcal{P}) \text{ para toda } m \in M \cap T\}.$$

Nótese además que

$$\mathcal{Q}[i] \subseteq \mathcal{R}[i] \subseteq \mathcal{P}[i] \text{ para todo } i \in H \cup M,$$

y las preferencias coinciden en cada caso.

Sea μ el emparejamiento hombre-óptimo para $\mathcal{R}$. Por la construcción de $\mathcal{R}$, $\mu(m) \succ_m \mathrm{PPF}(m, \mathcal{P})$ en $\mathcal{P}$ para toda $m \in M \cap T$.

Para terminar la prueba, se verificará lo mismo para $H \cap T$. Sea ν el emparejamiento mujer-óptimo para $\mathcal{Q}$. Entonces, por la construcción de $\mathcal{Q}$,

$$\nu(i) \succ_i \mathrm{PPF}(i, \mathcal{P}) \text{ para todo } i \in T.$$

A modo de contradicción supóngase que $h \in T$ es el primer hombre en ser rechazado por su respectiva $\nu(h)$ en el desarrollo del algoritmo de Gale–Shapley *hombres-proponen* en $\mathcal{R}$. Ahora bien, $\nu(h)$ sólo rechaza a h si tiene una propuesta mejor de otro hombre h' tal que

$$h' \succ_m h \text{ en } \mathcal{R}.$$

Como h es el primer hombre en ser rechazado por $\nu(h)$,

$$m \succ_{h'} \nu(h') \text{ en } \mathcal{R}.$$

Estas preferencias se cumplen también en $\mathcal{Q}$, y por tanto h' y m $\mathcal{Q}$-bloquean el emparejamiento $\mathcal{Q}$-estable ν, lo cual es imposible. Se concluye que ningún $h \in H \cap T$ es rechazado por $\nu(h)$ en el algoritmo para μ, de forma que

$$\mu(h) \succeq_h \nu(h) \succ_h \mathrm{PPF}(h, \mathcal{P}) \text{ para todo } h \in H \cap T,$$

como se quería demostrar. □

Teorema 4.17. Sea $T \subseteq H \cup M$ como en el enunciado del teorema 4.16, y sea $\mathcal{S}$ la lista de preferencias obtenida a partir de $\mathcal{P}$ permitiendo que cada $i \in T$ elimine a $\mathrm{PPF}(i, \mathcal{P})$. Entonces

$$\mathrm{Est}(\mathcal{S}) = \{\mu \in \mathrm{Est}(\mathcal{P}) \mid \mu(i) \neq \mathrm{PPF}(i, \mathcal{P}) \text{ para todo } i \in T\}.$$

Demostración. Por el teorema 4.16, existe $\mu \in \mathrm{Est}(\mathcal{P})$ tal que $\mu(i) \neq \mathrm{PPF}(i, \mathcal{P})$ para todo $i \in T$. Escríbase $T = \{i_1, \ldots, i_N\}$ y obténganse listas

$$\mathcal{P}^1, \mathcal{P}^2, \ldots, \mathcal{P}^N = \mathcal{S}$$

tales que en $\mathcal{P}^K$, i_K descalifica a $\mathrm{PPF}(i_K, \mathcal{P})$. Por el teorema 4.11, en cada paso μ sigue siendo estable, y

$$\begin{aligned}
\mathrm{Est}(\mathcal{S}) &= \{\mu \in \mathrm{Est}(\mathcal{P}^{N-1}) \mid \mu(i_N) \neq \mathrm{PPF}(i_N, \mathcal{P})\} \\
&= \{\mu \in \mathrm{Est}(\mathcal{P}^{N-2}) \mid \mu(i_K) \neq \mathrm{PPF}(i_K, \mathcal{P}) \text{ para } K = N, N-1\} \\
&\ \vdots \\
&= \{\mu \in \mathrm{Est}(\mathcal{P}^{1}) \mid \mu(i_K) \neq \mathrm{PPF}(i_K, \mathcal{P}) \text{ para } 2 \leq K \leq N\} \\
&= \{\mu \in \mathrm{Est}(\mathcal{P}) \mid \mu(i) \neq \mathrm{PPF}(i, \mathcal{P}) \text{ para todo } i \in T\}.
\end{aligned}$$ □

Este resultado es la última pieza en la formulación del algoritmo de emparejamiento multivaluado, que se enuncia a continuación.

4.3.2. Formulación del algoritmo

Dados dos conjuntos finitos no vacíos H y M y una lista de preferencias $\mathcal{P}$ entre H y M, se genera una lista de preferencias $\mathcal{P}^{**}$ según el siguiente procedimiento.

1. Sea $\mathcal{P}^*$ la lista que resulta de aplicar el algoritmo de Gale–Shapley bilateral a $\mathcal{P}$.

2. Si $|\mathcal{P}^*[i]| \leq 2$ para todo $i \in H \cup M$, se escoge $\mathcal{P}^{**} = \mathcal{P}^*$ y el algoritmo termina.

3. De lo contrario, sea $\mathcal{Q}$ la lista de preferencias obtenida a partir de $\mathcal{P}^*$ permitiendo que cada $i \in H \cup M$ elimine a $\mathrm{PPF}(i, \mathcal{P}^*)$.

4. Aplicando el algoritmo de Gale–Shapley bilateral a $\mathcal{Q}$, se calcula el conjunto $T = \{i \in H \cup M \mid i \text{ no es soltero en } \mathcal{Q}\}$.

5. Sea ahora $\mathcal{P}$ la lista de preferencias obtenida a partir de $\mathcal{P}^*$ permitiendo que los individuos de T eliminen a su peor pareja factible, y vuélvase al paso 1.

Se tienen los siguientes hechos:

- En el paso 1, $\mathrm{Est}(\mathcal{P}^*) = \mathrm{Est}(\mathcal{P})$, como se vio en la subsección 4.2.2. Además, si λ es un emparejamiento multivaluado $\mathcal{P}^*$-estable, entonces es $\mathcal{P}$-estable, como se demostrará más adelante (teorema 4.18).

- En el paso 2, si el algoritmo termina, entonces $\mathcal{P}^*$ define un emparejamiento multivaluado $\mathcal{P}$-estable según la construcción del teorema 4.15.

- En el paso 4, $T \neq \varnothing$, como se demostrará posteriormente (teorema 4.19). Esto implica que el algoritmo termina tras un número finito de pasos, pues en cada iteración la lista de preferencias es estrictamente más pequeña.

- En el paso 5, $\mathrm{Est}(\mathcal{P}) = \mathrm{Est}(\mathcal{P}^*)$. Más aún, si λ es un emparejamiento multivaluado $\mathcal{P}$-estable, entonces es $\mathcal{P}^*$-estable, como se demostrará a continuación (teorema 4.20).

- El algoritmo satisface indiferencia de género e indiferencia entre pares, ya que las operaciones hechas en cada paso no requieren distinguir entre géneros ni entre individuos.

Estas observaciones se combinan para demostrar que el algoritmo termina, produce un emparejamiento multivaluado $\mathcal{P}$-estable, y posee indiferencia de género e indiferencia entre pares. Para terminar esta prueba, se verifican los hechos pendientes de demostración.

Teorema 4.18. En el paso 1, si λ es un emparejamiento multivaluado $\mathcal{P}^*$-estable, entonces es $\mathcal{P}$-estable.

Demostración. Por la $\mathcal{P}^*$-estabilidad de λ, para todo individuo i y todo $k \in \lambda(i)$,

$$k \succeq_i \mathrm{PPF}(i, \mathcal{P}).$$

Supóngase a modo de contradicción que λ es $\mathcal{P}$-bloqueado. Sean $i, j, i_\lambda \in H \cup M$ tales que

$$j \succ_i i_\lambda \in \lambda(i) \text{ en } \mathcal{P},$$
$$i \succ_j j_\lambda \in \lambda(j) \text{ en } \mathcal{P}.$$

El teorema 4.13 demuestra que $j \in \mathcal{P}^*[i]$ e $i \in \mathcal{P}^*[j]$, y por tanto estas preferencias también valen en $\mathcal{P}^*$. Se sigue que i y j $\mathcal{P}^*$-bloquean λ, contradiciendo que λ es $\mathcal{P}^*$-estable. $\square$

Teorema 4.19. En el paso 4, $T \neq \varnothing$.

Demostración. Dado que el algoritmo no terminó en el paso 2, por la construcción de $\mathcal{P}^*$ existen $h \in H$, $m \in M$ tales que

$$\mathrm{MPF}(h, \mathcal{P}) \succ_h m \succ_h \mathrm{PPF}(h, \mathcal{P}) \text{ en } \mathcal{P}^*,$$
$$\mathrm{MPF}(m, \mathcal{P}) \succ_m h \succ_m \mathrm{PPF}(m, \mathcal{P}) \text{ en } \mathcal{P}^*.$$

Por tanto, $h \in \mathcal{Q}[m]$ y $m \in \mathcal{Q}[h]$. En consecuencia, si ν es un emparejamiento $\mathcal{Q}$-estable, entonces no pueden estar solteros simultáneamente h y m. Se concluye que $h \in T$ o $m \in T$. $\square$

Teorema 4.20. En el paso 5, si λ es un emparejamiento multivaluado $\mathcal{P}$-estable, entonces es $\mathcal{P}^*$-estable.

Demostración. Como λ es $\mathcal{P}$-estable, para cada $i \in H \cup M$ y $k \in \lambda(i)$ se tiene

$$k \succeq_i \mathrm{PPF}(i, \mathcal{P}).$$

Supóngase a modo de contradicción que λ es $\mathcal{P}^*$-bloqueado. Sean $i, j, i_\lambda \in H \cup M$ tales que

$$j \succ_i i_\lambda \in \lambda(i) \text{ en } \mathcal{P}^*,$$
$$i \succ_j j_\lambda \in \lambda(j) \text{ en } \mathcal{P}^*.$$

Por la construcción de $\mathcal{P}$, estas preferencias también se tienen en $\mathcal{P}$. Se sigue que i y j $\mathcal{P}$-bloquean λ, contradiciendo su $\mathcal{P}$-estabilidad. $\square$

4.3.3. Un ejemplo

Considérese la siguiente lista de preferencias $\mathcal{P}$ entre seis hombres $H = \{\alpha, \beta, \ldots, \zeta\}$ y seis mujeres $M = \{A, B, \ldots, F\}$:

α	β	γ	δ	ε	ζ	A	B	C	D	E	F
E	A	A	B	C	D	δ	γ	ζ	ε	β	α
C	F	E	F	A	B	ζ	ε	β	β	δ	γ
A	B	C	D	E	F	α	α	δ	γ	ζ	ε
B	C	D	C	F	E	ε	β	γ	δ	ε	ζ
D	D	F	E	B	A	γ	ζ	α	α	γ	δ
F	E	B	A	D	C	β	δ	ε	ζ	α	β

El primer paso es aplicar el algoritmo de Gale–Shapley bilateral. El único rechazo que se hace antes de terminar el proceso es el rechazo de β por parte de A, lo cual produce la siguiente lista $\mathcal{P}^*$:

α	β	γ	δ	ε	ζ	A	B	C	D	E	F
E		A	B	C	D	δ	γ	ζ	ε	β	α
C	F	E	F	A	B	ζ	ε	β	β	δ	γ
A	B	C	D	E	F	α	α	δ	γ	ζ	ε
B	C	D	C	F	E	ε	β	γ	δ	ε	ζ
D	D	F	E	B	A	γ	ζ	α	α	γ	δ
F	E	B	A	D	C		δ	ε	ζ	α	β

Puesto que la lista de cada individuo todavía tiene más de dos parejas, el proceso continúa. Se obtiene la lista $\mathcal{Q}$ eliminando la última pareja de cada individuo:

α	β	γ	δ	ε	ζ	A	B	C	D	E	F
E		A	B	C	D	δ	γ	ζ	ε	β	α
C	F	E	F	A	B	ζ	ε	β	β	δ	γ
A	B	C	D	E	F	α	α	δ	γ	ζ	ε
B	C	D	C	F	E	ε	β	γ	δ	ε	ζ
D	D	F	E	B	A		ζ	α	α	γ	δ

En el paso 4, aplicando el algoritmo de Gale–Shapley bilateral a $\mathcal{Q}$ se obtiene la siguiente sucesión de rechazos:

todos son rechazados por su primera pareja,
β rechaza a D y B rechaza a ζ,
γ rechaza a F y F rechaza a δ,
ε rechaza a B y D rechaza a α.

El resultado es la siguiente lista, que se denotará por $\mathcal{R}$:

α	β	γ	δ	ε	ζ	A	B	C	D	E	F
C		E		A		ζ		β		δ	
A	B	C	D	E	F	α	α	δ	γ	ζ	ε
B	C	D	C	F	E	ε	β	γ	δ	ε	ζ
			E		A			α		γ	

Puesto que nadie queda soltero en $\mathcal{R}$, en el paso 5 se asignará $\mathcal{P} = \mathcal{Q}$, y al regresar al paso 1 se obtiene $\mathcal{P}^* = \mathcal{R}$.

Puesto que todavía hay individuos con más de dos parejas en lista, se procede al paso 3. Se obtiene un nuevo $\mathcal{Q}$ eliminando la última pareja de cada persona:

α	β	γ	δ	ε	ζ	A	B	C	D	E	F
C		E		A		ζ		β		δ	
A	B	C	D	E	F	α	α	δ	γ	ζ	ε
			C		E			γ		ε	

Al ejecutar el algoritmo de Gale–Shapley bilateral sobre el nuevo $\mathcal{Q}$ se obtienen los siguientes rechazos:

todos son rechazados por su primera pareja,
C rechaza a γ y E rechaza a ε.

El resultado es esta lista,

α	β	γ	δ	ε	ζ	A	B	C	D	E	F
A			C		E	α		δ		ζ	

en la cual los no solteros son $T = \{\alpha, \delta, \zeta, A, C, E\}$. En el paso 5, se obtiene el perfil $\mathcal{P}$ permitiendo que estas seis personas eliminen a su última pareja:

α	β	γ	δ	ε	ζ	A	B	C	D	E	F
C		E		A		ζ		β		δ	
A	B	C	D	E	F	α	α	δ	γ	ζ	ε
	C	D	C	F	E		β	γ	δ	ε	ζ

Volviendo al paso 1, se ejecuta el algoritmo de Gale–Shapley bilateral para este nuevo $\mathcal{P}$, obteniendo los siguientes rechazos:

$$\alpha, \delta, \zeta, A, C, E \text{ rechazan respectivamente a } B, E, A, \varepsilon, \alpha, \gamma,$$
$$\beta \text{ rechaza a } C,$$

y el resultado es el siguiente:

α	β	γ	δ	ε	ζ	A	B	C	D	E	F
A	B	C	D	E	F	α		δ	γ	ζ	ε
		D	C	F	E		β	γ	δ	ε	ζ

Puesto que todas las listas han sido reducidas a dos personas o menos, se ha encontrado $\mathcal{P}^{**}$ y el algoritmo termina.

Obsérvese que existen cuatro emparejamientos univaluados μ contenidos en esta lista de preferencias, uno por cada elección de $\mu(\gamma) \in \{C, D\}$ y $\mu(\varepsilon) \in \{E, F\}$, y todos ellos son estables por el teorema 4.7.

4.3.4. Consecuencias

Con la formulación de este algoritmo, se ha comprobado que existe una generalización del concepto de estabilidad en emparejamientos de una forma que, a diferencia de la estabilidad clásica, es compatible al mismo tiempo con la indiferencia de género y la indiferencia entre pares. Merecen mencionarse, sin embargo, varias dificultades que persisten aun formulado el algoritmo.

Para empezar, si bien el emparejamiento multivaluado λ que resulta de aplicar el algoritmo siempre puede escribirse como unión de a lo más dos emparejamientos univaluados estables, en todo caso puede contener un número arbitrariamente grande de subemparejamientos univaluados. Por ejemplo, considérese el perfil de preferencias entre los conjuntos

$$H_n = \{\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_n, \beta_n\}, \qquad M_n = \{A_1, B_1, A_2, B_2, \ldots, A_n, B_n\}$$

dado por

α_r	β_r	A_r	B_r
A_r	B_r	β_r	α_r
B_r	A_r	α_r	β_r
$\vdots$	$\vdots$	$\vdots$	$\vdots$

para $r = 1, \ldots, n$,

donde los puntos suspensivos indican el resto de las posibles parejas. Este perfil puede considerarse una combinación de n veces el perfil de la página 23, y por tanto tiene 2^n emparejamientos estables μ, uno por cada elección de $\mu(\alpha_r) \in \{A_r, B_r\}$ para $r = 1, \ldots, n$. Al ejecutar el algoritmo sobre este perfil, las listas de preferencias quedan reducidas a dos personas en el primer paso, y por tanto λ contiene todos los emparejamientos univaluados estables.

Así pues, si el objetivo es fijar un único emparejamiento univaluado contenido en λ, esto puede resultar en general difícil, y en particular no puede hacerse sin sacrificar la indiferencia de género o la indiferencia entre pares.

Otra dificultad es que no es claro el comportamiento del algoritmo con respecto a las medidas de inequidad de la sección 3.2. El algoritmo produce $\mu_1 \cup \mu_2$ en el ejemplo de la subsección 3.2.1, μ_3 en el de la subsección 3.2.2, y μ_H en el de la subsección 3.2.3. Puede verse, por tanto, que el algoritmo no minimiza de manera consistente ninguna de las tres medidas de inequidad. A pesar de esto, el algoritmo logra en general (salvo en el ejemplo de la subsección 3.2.1) identificar el emparejamiento intuitivamente más equitativo.

Capítulo 5

Conclusiones

A lo largo de este trabajo se ha visto que el problema de la equidad de género en emparejamientos estables no es de fácil abordaje. El teorema 3.5 en particular plantea una dificultad esencial: para obtener equidad de género es preciso discriminar entre individuos. El concepto de solución y el algoritmo presentados en la sección 4 pretenden ofrecer un camino alterno que sí es compatible con ambas nociones de justicia. Si bien la solución propuesta excede el ámbito del problema original del emparejamiento estable, ofrece una dirección en la cual puede generalizarse el problema para obtener la equidad buscada. Ya sabiendo que la solución con equidad de género existe siempre en emparejamientos multivaluados, una posible área de trabajo futuro consiste en determinar si es posible restringir el concepto de solución sin sacrificar ninguna de estas nociones de justicia.

Los ejemplos vistos en la sección 2 dan a entender que las medidas de inequidad discutidas no logran fijar adecuadamente la idea de equidad en emparejamientos estables: el esfuerzo por minimizar una de estas medidas no siempre logra dar con un emparejamiento equitativo y puede incluso ser contraproducente. A pesar de esto, las medidas de inequidad sin duda son útiles como principios generales, y puede pensarse en emplear una combinación de ellas como heurística para encontrar emparejamientos equitativos. La interacción de estas medidas con los resultados del algoritmo también merece ser estudiada; pueden tratar de definirse extensiones de estas medidas a emparejamientos multivaluados y buscar maneras de minimizarlas sin sacrificar la indiferencia de género o entre pares.

En ese orden de ideas, cabe observar que este algoritmo es apenas uno de posiblemente muchos que satisface las nociones de justicia. En particular, las eliminaciones del paso 5 pueden realizarse con un criterio distinto; otra línea de trabajo futuro consiste en formular otros algoritmos de emparejamiento multivaluado que tengan propiedades diferentes y conserven las nociones de justicia. Alguno de estos algoritmos podría emplear la estructura combinatoria del conjunto de emparejamientos estables tal como lo hace el algoritmo mencionado en la subsección 3.2.2, para lo cual haría falta estudiar cómo encajan los emparejamientos multivaluados estables en la estructura de retículo de los emparejamientos univaluados. Otra vía de trabajo consiste en identificar propiedades que caractericen de manera única alguna solución multivaluada particular, tal y como la solución de Gale–Shapley queda caracterizada por la estabilidad y optimalidad para los hombres (o para las mujeres, según el caso).

Cabe mencionar, por último, que el algoritmo presentado en este trabajo no es una salida definitiva a la imposibilidad de la equidad de género: en el ejemplo de dos hombres y dos mujeres de la página 23, es definitivamente imposible escoger un emparejamiento univaluado sin favorecer a uno de los géneros sobre el otro. Esto sugiere que, en cierto sentido, la "poligamia" es esencialmente la solución equitativa, al menos en algunos casos. La existencia de un algoritmo multivaluado equitativo le presta justificación a la idea de que la bigamia es, como mínimo, digna de consideración. Sin embargo, deben reconocerse las limitaciones prácticas del modelo multivaluado, que obran en contra de esta conclusión; por ejemplo, el modelo univaluado admite resultados naturales sobre lo que ocurre cuando algún individuo entra o sale del sistema [16, 20]. No es claro si este tipo de dinámica se extiende al modelo multivaluado de manera igualmente natural. El modelo tampoco refleja la existencia de emparejamientos múltiples donde una persona puede tener tres o más parejas, lo cual también ocurre en la práctica.

En síntesis, los resultados presentados en este trabajo son apenas una garantía de existencia para la equidad de género en emparejamientos estables. Esta modesta contribución a la teoría no exime de los trabajos posteriores necesarios para determinar el verdadero alcance de este concepto de solución.

Bibliografía

[1] David Gale y Lloyd S. Shapley. “College admissions and the stability of marriage”. En: *The American Mathematical Monthly* 69.1 (1962), págs. 9-15.

[2] Donald Ervin Knuth. *Mariages stables et leurs relations avec d’autres problèmes combinatoires: introduction à l’analyse mathématique des algorithmes*. Montréal: Presses de l’Université de Montréal, 1976.

[3] Alvin E Roth. “The economics of matching: Stability and incentives”. En: *Mathematics of operations research* 7.4 (1982), págs. 617-628.

[4] Charles Blair. “Every finite distributive lattice is a set of stable matchings”. En: *Journal of Combinatorial Theory, Series A* 37.3 (1984), págs. 353-356.

[5] David Gale y Marilda Sotomayor. “Ms. Machiavelli and the stable matching problem”. En: *The American Mathematical Monthly* 92.4 (1985), págs. 261-268.

[6] David Gale y Marilda Sotomayor. “Some remarks on the stable matching problem”. En: *Discrete Applied Mathematics* 11.3 (1985), págs. 223-232.

[7] Robert W Irving. “An efficient algorithm for the “stable roommates” problem”. En: *Journal of Algorithms* 6.4 (1985), págs. 577-595.

[8] Alvin E Roth. “Conflict and coincidence of interest in job matching: some new results and open questions”. En: *Mathematics of Operations Research* 10.3 (1985), págs. 379-389.

[9] Alvin E Roth. “The college admissions problem is not equivalent to the marriage problem”. En: *Journal of economic Theory* 36.2 (1985), págs. 277-288.

[10] Dan Gusfield. “Three fast algorithms for four problems in stable marriage”. En: *SIAM Journal on Computing* 16.1 (1987), págs. 111-128.

[11] Robert W Irving, Paul Leather y Dan Gusfield. “An efficient algorithm for the “optimal” stable marriage”. En: *Journal of the ACM (JACM)* 34.3 (1987), págs. 532-543.

[12] Charles Blair. “The lattice structure of the set of stable matchings with multiple partners”. En: *Mathematics of operations research* 13.4 (1988), págs. 619-628.

[13] F Masarani y Sadik S Gokturk. “On the existence of fair matching algorithms”. En: *Theory and Decision* 26.3 (1989), págs. 305-322.

[14] Alvin E Roth y Marilda Sotomayor. “Two-sided matching”. En: *Handbook of game theory with economic applications* 1 (1992), págs. 485-541.

[15] Robert W Irving. “Stable marriage and indifference”. En: *Discrete Applied Mathematics* 48.3 (1994), págs. 261-272.

[16] Yosef Blum, Alvin E Roth y Uriel G Rothblum. “Vacancy chains and equilibration in senior-level labor markets”. En: *Journal of Economic theory* 76.2 (1997), págs. 362-411.

[17] Marilda Sotomayor. “Three remarks on the many-to-many stable matching problem”. En: *Mathematical social sciences* 38.1 (1999), págs. 55-70.

[18] Mourad Baıou y Michel Balinski. “Many-to-many matching: stable polyandrous polygamy (or polygamous polyandry)”. En: *Discrete Applied Mathematics* 101.1-3 (2000), págs. 1-12.

[19] Atila Abdulkadiroğlu y Tayfun Sönmez. “School choice: A mechanism design approach”. En: *American economic review* 93.3 (2003), págs. 729-747.

[20] Péter Biró, Katarína Cechlárová y Tamás Fleiner. “The dynamics of stable matchings and half-matchings for the stable marriage and roommates problems”. En: *International Journal of Game Theory* 36.3-4 (2008), págs. 333-352.

[21] Maria Silvia Pini y col. “Manipulation complexity and gender neutrality in stable marriage procedures”. En: *Autonomous Agents and Multi-Agent Systems* 22.1 (2011), págs. 183-199.

[22] Maria Silvia Pini y col. “Weights in stable marriage problems increase manipulation opportunities”. En: *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*. ACM. 2011, págs. 200-204.

[23] Atila Abdulkadiroğlu y Tayfun Sönmez. “Matching markets: Theory and practice”. En: *Advances in Economics and Econometrics* 1 (2013), págs. 3-47.

[24] Ioannis Giannakopoulos y col. “An equitable solution to the stable marriage problem”. En: *Tools with Artificial Intelligence (ICTAI), 2015 IEEE 27th International Conference on*. IEEE. 2015, págs. 989-996.

Índice de definiciones y teoremas

Índice alfabético