KAPITEL 1 — EINFÜHRUNG: STABILE MATCHINGS

F. VALLENTIN, A. GUNDERT

In diesem Kapitel werden wir ein erstes konkretes Problem des Operations Research kennenlernen. Es handelt sich um das Problem des stabilen Matchings, ein wichtiges Zuordnungsproblem, das viele Anwendungen in der nichtmonetären Ökonomie besitzt. Im Jahr 2012 bekamen Lloyd S. Shapley and Alvin E. Roth den Nobelpreis für Ökonomie für *the theory of stable allocations and the practice of market design*.

1. Grundbegriffe

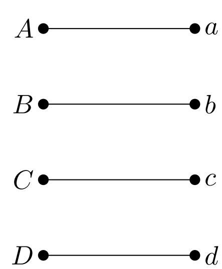
Definition 1.1. *Gegeben seien Mengen*

$$\mathcal{M} = \{A_1, \dots, A_n\} \quad \textit{und} \quad \mathcal{F} = \{B_1, \dots, B_n\}$$

mit je n Elementen. Ein Matching *ist eine Bijektion $\sigma\colon \mathcal{M} \to \mathcal{F}$.*

Dieses hat die folgende Interpretation[1]: Die n-elementige Menge $\mathcal{M}$ stellt n Männer dar, die n-elementige Menge $\mathcal{F}$ n Frauen. Ein Matching $\sigma\colon \mathcal{M} \to \mathcal{F}$ entspricht n „traditionellen“ Hochzeiten: Es handelt sich um monogame, heterosexuelle Ehen, wobei niemand ledig ist.
Desweiteren stellen wir uns vor, dass jeder Mann eine *Präferenzliste* für die n Frauen besitzt und umgekehrt jede Frau eine Präferenzliste für die n Männer besitzt. Beide Präferenzlisten werden durch zwei Matrizen mit jeweils n^2 Einträgen dargestellt.

Beispiel 1.2. *Hier ist ein einfaches Beispiel für ein Matching mit vier Männernn $\mathcal{M} = \{A, B, C, D\}$ und vier Frauen $\mathcal{F} = \{a, b, c, d\}$.*



Folgende Tabellen sind Beispiele für Präferenzlisten:

Männer				
$A:$	c	b	d	a
$B:$	b	a	c	d
$C:$	b	d	a	c
$D:$	c	a	d	b

Frauen				
$a:$	A	B	D	C
$b:$	C	A	D	B
$c:$	C	B	D	A
$d:$	B	A	C	D

Date: 28. April 2014.

[1] Den anachronistischen Charakter dieses Modells nehmen wir unkommentiert hin.

Dabei ist die Präferenz höher, je weiter links eine Person gelistet ist. Zum Beispiel bevorzugt Mann A Frau c gegenüber Frau b, Frau b gegenüber Frau d und diese wiederum gegenüber Frau a.

Notation 1.3. *Wir schreiben zum Beispiel*

$$c >_A b \iff \text{Mann } A \text{ bevorzugt Frau } c \text{ gegenüber Frau } b$$

und analog

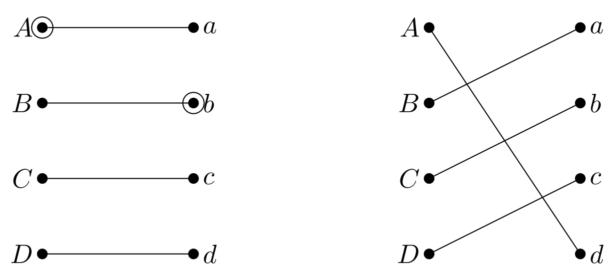
$$A >_a B \iff \text{Frau } a \text{ bevorzugt Mann } A \text{ gegenüber Mann } B.$$

Definition 1.4. *Ein Matching $\sigma\colon \mathcal{M} \to \mathcal{F}$ heißt* stabil*, wenn für alle $M \in \mathcal{M}$ und alle $f \in \mathcal{F}$ mit $\sigma(M) \neq f$ stets*

$$\sigma(M) >_M f \quad \text{oder} \quad \sigma^{-1}(f) >_f M$$

gilt. Ansonsten könnte Interesse an einem Seitensprung bestehen.

In unserem Beispiel:



ist nicht stabil ist stabil

Die Bestimmung von stabilen Matchings hat viele Anwendungen bei Zuordnungsproblem, z.B. wird seit 1952 in den USA im Rahmen des National Resident Matching Program Medizinstudenten auf Krankenhäuser verteilt, wobei ein Verfahren angewendet wird, das auf dem Konzept der stabilen Matchings beruht. In realistischen Situationen muss man das Modell anpassen, da z.B. die Präferenztabellen nicht vollständig sind, oder die Mengen $\mathcal{M}$ und $\mathcal{F}$ unterschiedliche Kardinalität aufweisen.

Die Definition wirft zwei Fragen auf:

(1) Wie findet man ein stabiles Matching?
(2) Gibt es immer eins?

Die erste Frage kann man naiv beantworten: Teste einfach alle $n!$ Möglichkeiten. Das ist aber keine gute Idee, weil die Rechenzeit schnell sehr groß wird, wie die folgende Tabelle zeigt:

n	Rechenzeit (Annahme: pro Bijektion 10^{-9} Sekunden)
10	$0,03$ Sekunden
15	21 Minuten
20	77 Jahre
25	$4,9 \cdot 10^8$ Jahre ($\approx \frac{1}{10}$ Alter der Erde)

2. Gale-Shapley Algorithmus

Der Algorithmus von Gale und Shapley gibt eine deutlich bessere Antwort auf die erste Frage. Außerdem kann mit ihm auch die zweite Frage positiv beantworten.

Algorithmus 2.1. *(in Pseudocode)*

Verlobe Frauen $a_1, \ldots, a_n$ *mit virtuellem Mann* A_0*, den jede Frau am wenigsten bevorzugt. ;*
for $k = 1$ ***to*** n **do**
    $M = A_k$
    while $M \neq A_0$ **do**
        f = *beste verbleibende Frau auf* M*'s Liste*
        M *macht* f *Heiratsantrag*
        if $M >_f$ *aktueller Verlobter von* f **then**
            verlobe M *und* f
            M = *bisheriger Verlobter von* f
        end
        if $M \neq A_0$ **then**
            streiche f *von* M*'s Liste*
        end
    end
end

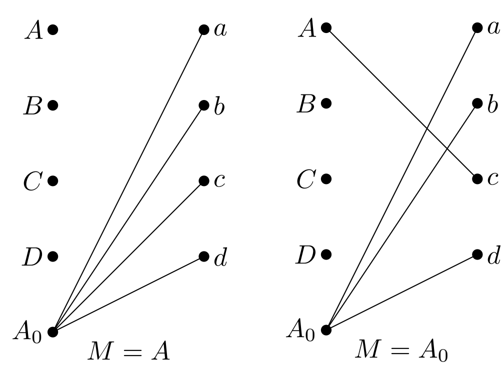
Algorithmus 1 : Gale-Shapley Algorithmus

In unserem Beispiel:

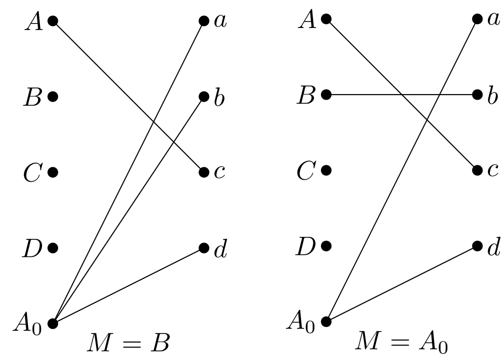
Männer				
$A:$	$\cancel{c}$	$\cancel{b}$	d	a
$B:$	$\cancel{b}$	a	c	d
$C:$	b	d	a	c
$D:$	c	a	d	b

Frauen					
$a:$	A	B	D	C	A_0
$b:$	C	A	D	B	A_0
$c:$	C	B	D	A	A_0
$d:$	B	A	C	D	A_0

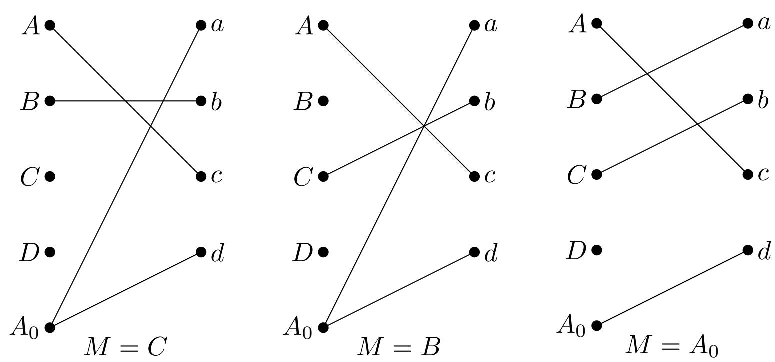
$\underline{k = 1}$:



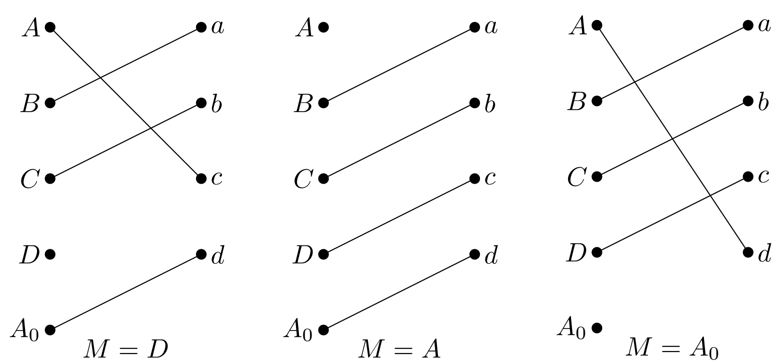
<u>$k = 2$</u>:



<u>$k = 3$</u>:



<u>$k = 4$</u>:



Das letzte Matching ist stabil, was auch folgender Satz zeigt.

Satz 2.2. *Der Gale-Shapley Algorithmus berechnet ein stabiles Matching. Insbesondere gibt es immer eines.*

Beweis.

(1) Der Algorithmus ist wohldefiniert: Die Listen der Männer sind niemals leer. Denn angenommen die Liste von M wäre leer. Dann ist M von allen Frauen abgewiesen worden. Jede Frau ist aber höchstens mit einem Mann verlobt, die Situation der Frauen verschlechtert sich nie und M ist stets besser als A_0. Es muss also neben M (und A_0) n „echte “Männer geben. Widerspruch

(2) Das resultierende Matching ist stabil. Denn angenommen $\sigma(M) \neq f$ und $f >_M \sigma(M)$. Dann ist M im Laufe des Algorithmus von f abgewiesen worden und $\sigma^{-1}(f) >_f M$. Somit ist die Stabilität nicht verletzt.

$\square$

Bemerkung 2.3. *Im Algorithmus werden (für $n \geq 2$) weniger als n^2 Heiratsanträge gemacht (siehe erstes Aufgabenblatt) $\rightarrow$ Viel effizienter als $n!$.*

Satz 2.4. *Der Gale-Shapley Algorithmus berechnet das für die Männer beste stabile Matching, das heißt wenn f von M's Liste entfernt wird, dann gibt es kein stabiles Matching σ mit $\sigma(M) = f$.*

Für den Beweis benötigen wir noch das folgende Lemma:

Lemma 2.5. *Falls „streiche f von M's Liste "aufgerufen wird, gibt es einen Mann $\tilde{M}$ mit $\tilde{M} >_f M$ und $f >_{\tilde{M}}$ alle Frauen auf $\tilde{M}$'s verbleibenden Liste.*

Beweis. „Streiche f von M's Liste "passiert, wenn

(1) M möchte f heiraten, aber es gibt einen Mann $\tilde{M}$, den aktuellen Verlobten von f mit $\tilde{M} >_f M$.
(2) M und f waren verlobt, aber es gibt einen Mann $\tilde{M}$ mit $\tilde{M} >_f M$, der f einen Heiratsantrag gemacht hat.

Es gibt also in beiden Situationen einen Mann $\tilde{M}$ mit $\tilde{M} >_f M$ für den auch gilt: $f >_{\tilde{M}}$ alle Frauen auf $\tilde{M}$'s verbleibender Liste. $\square$

Beweis. (des Satzes): Wir zeigen: Für alle stabilen Matchings σ und $M \in \mathcal{M}, f \in \mathcal{F}$ gilt: $\sigma(M) = f \quad \Rightarrow f$ wird nicht von M's Liste gestrichen.
Angenommen das stimmt nicht. Dann gibt es ein stabiles Matching σ mit mindestens einem Paar (M, f) sodass

- $\sigma(M) = f$ und
- f wird von M's Liste gestrichen.

Wähle unter diesen Paaren (M, f) dasjenige aus, bei dem die „Streichung "am frühesten passiert (im Laufe des Algorithmus). Nach dem Lemma gibt es $\tilde{M}$ mit $\tilde{M} >_f M$ und $f >_{\tilde{M}}$ alle Frauen auf $\tilde{M}$'s verbleibender Liste (zum Zeitpunkt der Streichung). Betrachte $\sigma(\tilde{M})$: Da σ stabil, gilt $\sigma(\tilde{M}) >_{\tilde{M}} f = \sigma(M)$. Das heißt $\sigma(\tilde{M})$ wurde zu einem früheren Zeitpunkt von $\tilde{M}$'s Liste gestrichen. Widerspruch zur Wahl von (M, f). $\square$

Korollar 2.6. *Das stabile Matching, welches der Gale-Shapley Algorithmus berechnet, ist unabhängig von der Reihenfolge der Männer und Frauen.*

3. Geometrische Modellierung

Der Gale-Shapley Algorithmus ist ein *kombinatorischer Algorithmus* (verwendet keine Zahlen).

Vorteil: + sehr schnell, effizient

Nachteil: - relativ kompliziert
- sehr problemspezifisch (kann aber angepasst werden, z.B. für m Männer, n Frauen, unvollständige Präferenzlisten)

Jetzt: Geometrische Modellierung mit linearen Ungleichungen.

Vorteil: + einfache Modellierung
+ leicht veränderbar

Nachteil: - im Allgemeinen nicht die schnellste Lösungsmethode

Codiere ein Matching $\sigma\colon \mathcal{M} \to \mathcal{F}$, $\mathcal{M} = \{A_1, \ldots, A_n\}$, $\mathcal{F} = \{a_1, \ldots, a_n\}$ als Permutationsmatrix:

$$X \in \{0,1\}^{n\times n} \quad \text{mit} \quad X_{ij} = \begin{cases} 1, & \text{falls } \sigma(A_i) = a_j \\ 0, & \text{sonst.} \end{cases}$$

Satz 3.1. *$X \in \{0,1\}^{n\times n}$ codiert ein* stabiles *Matching $\Leftrightarrow$ X erfüllt folgende Ungleichungen:*

(1) $\sum_{j=1}^{n} X_{ij} \leq 1$ *für alle $i = 1, \ldots n$ (Zeilensummen)*

(2) $\sum_{i=1}^{n} X_{ij} \leq 1$ *für alle $j = 1, \ldots n$ (Spaltensummen)*

(3) $X_{ij} \geq 0$ *für alle $i, j = 1, \ldots n$*

(4) $X_{ij} + \sum_{k : a_k >_{A_i} a_j} X_{ik} + \sum_{k : A_k >_{a_j} A_i} X_{kj} \geq 1$ *für alle $i, j = 1, \ldots n$ (Stabilitätsbedingung)*

Beweis. Klar nach Definition von stabilem Matching. $\square$

Später in der Vorlesung:
Die Menge

$$\text{SMP} = \{X \in \mathbb{R}^{n\times n} : X \text{ erfüllt die Bedingungen (1)- (4) }\}$$

ist ein Polytop (Lösungsmenge eines Systems linearer Ungleichungen, die beschränkt ist). Zum Beispiel ein 5-Eck ist ein zweidimensionales Polytop oder ein Würfel ist ein dreidimensionales Polytop.

Vande Vate (1989):
Die Ecken des Polytops SMP codieren genau die stabilen Matchings.

Sei $C \in \mathbb{R}^{n\times n}$ eine Heiratsbewertungsmatrix. (C_{ij} gibt den Wert der Heirat zwischen A_i und a_j an). Dann erhält man das optimale stabile Matching bezüglich C mit dem linearen Optimierungsproblem:

$$\begin{aligned} \max \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij} \\ X \in \mathbb{R}^{n\times n} \\ X \text{ erfüllt Bedingungen (1)–(4)} \end{aligned}$$

Prof. Dr. F. Vallentin, Dr. A. Gundert, Mathematisches Institut, Universität zu Köln, Weyertal 86–90, 50931 Köln, Deutschland
E-mail address: frank.vallentin@uni-koeln.de, anna.gundert@uni-koeln.de