

22. Algorithmus der Woche

Partnerschaftsvermittlung

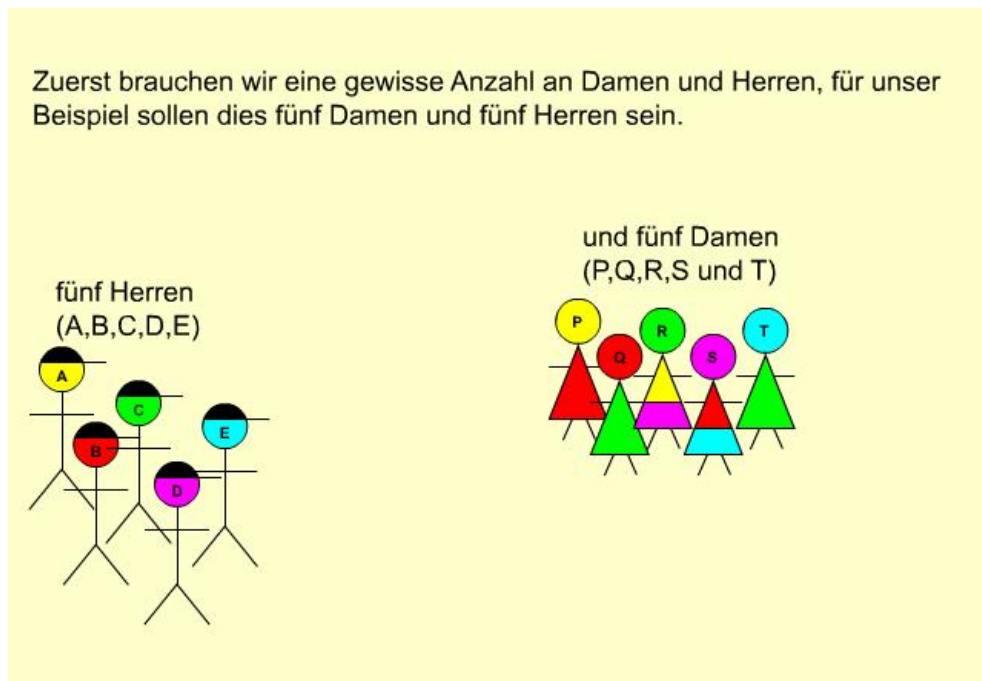
Drum prüfe, wer sich ewig bindet

Autor

Volker Claus, Universität Stuttgart
Volker Diekert, Universität Stuttgart
Holger Petersen, Universität Stuttgart

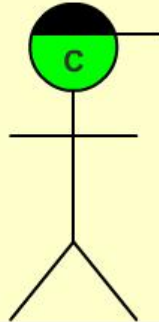
Eine Agentur für die Vermittlung von Partnerschaften versucht, möglichst viele Personen, die einander sympathisch sind, zusammenzubringen. Hierbei nehmen wir den Fall der klassischen Ehe an, d.h., es gibt eine Menge von Herren H und eine Menge von Damen D sowie eine Liste, die alle Kombinationen (h, d) von Herren h und Damen d umfasst, welche einander wechselseitig sympathisch finden und als Partner in Frage kommen.

In den Bildern weiter unten haben wir 5 Herren A, B, C, D, E und 5 Damen P, Q, R, S, T und eine zugehörige Liste aufgeführt. Die Agentur könnte in diesem Beispiel nacheinander die Paare (B, P) , (C, R) und (E, S) zusammenstellen; dann muss sie allerdings aufhören, weil nun A, D, Q und T keine Partner mehr bekommen können. Hätte sie dagegen (A, P) , (B, Q) , (C, R) , (D, S) und (E, T) gebildet, so hätte jede Person einen Partner erhalten (und die Agentur hätte mehr Geld für die Partnervermittlung verdient).



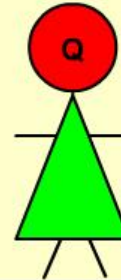
Schauen wir uns die Figuren, welche die Damen und Herren darstellen, genauer an.

Zuerst einen Herren:



Die Farben der Köpfe stellen hierbei die speziellen Vorzüge der Herren dar.

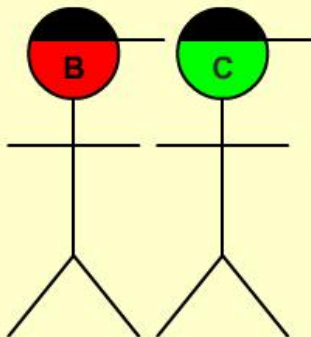
Und nun die Damen:



Bei den Damen stellen die Farben der Kleider und Köpfe deren Vorlieben dar.

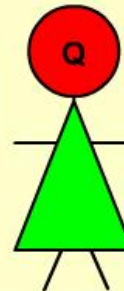
Damen haben also mehrere Favoriten (Herren, mit denen sich die Damen anfreunden wollen) dies bedeutete beispielsweise:

Den Herren B und den Herren C



Diese Dame interessiert sich also für zwei Herren

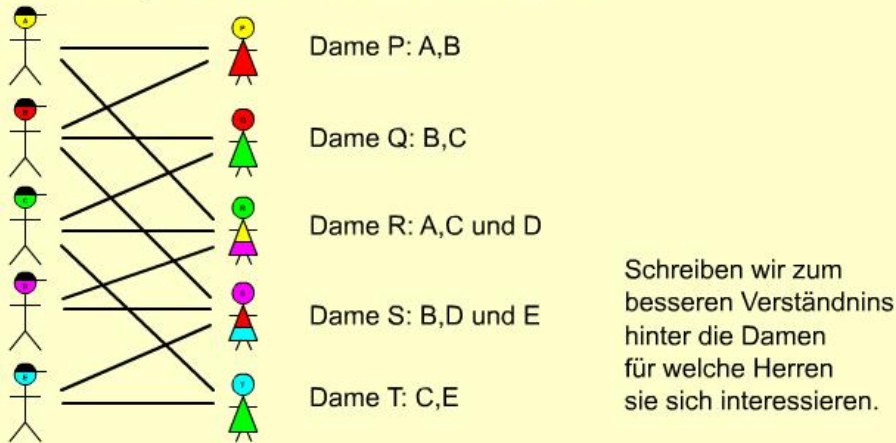
Die Dame Q:



roter Kopf
grünes Kleid

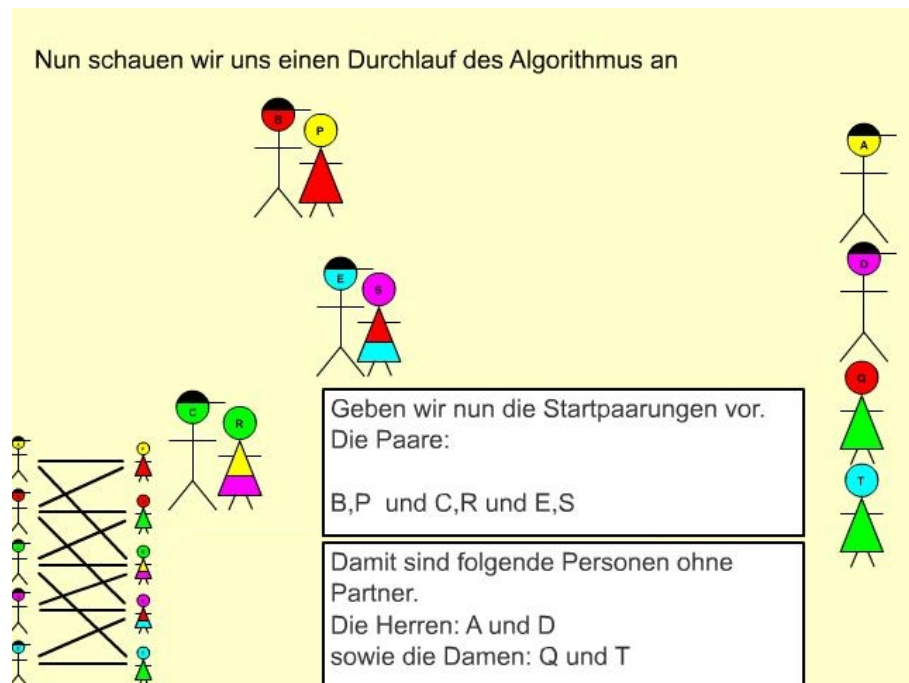
Dies ergibt eine Reihe von Beziehungen zwischen den Herren und den Damen. Diese kann man in Form eines sogenannten Graphen recht einfach darstellen.

Man verbindet die Herren und Damen, bei denen sich die Vorzüge, der Herren, mit den Vorlieben der Damen decken.



Es handelt sich bei der Partnervermittlung um ein Zuordnungsproblem: Finde zu zwei Mengen H und D und einer Liste von Paaren eine möglichst große Teilmenge dieser Paare, wobei jedes Element aus H und jedes Element aus D höchstens einmal auftreten darf. Dieses Problem ist unter dem Namen "Heiratsproblem" bekannt. Wir wollen hier einen Algorithmus angeben, der zu einer gegebenen Liste eine maximal große Teilmenge ermittelt.

Hierzu lassen wir die Damen mit ihren Herren tanzen, d.h., wir stellen nacheinander Paare zusammen, die in der Liste vorkommen. Die bereits gebildeten Paare nennen wir Tanzpaare und stellen uns vor, sie tanzen gemeinsam auf einer Tanzfläche, wobei die noch nicht vermittelten Personen am Rande der Fläche stehen. Beispielsweise werden in der Animation nacheinander die Paare (B,P) , (C,R) und (E,S) ausgewählt. Danach können wir kein weiteres Paar mehr auswählen.



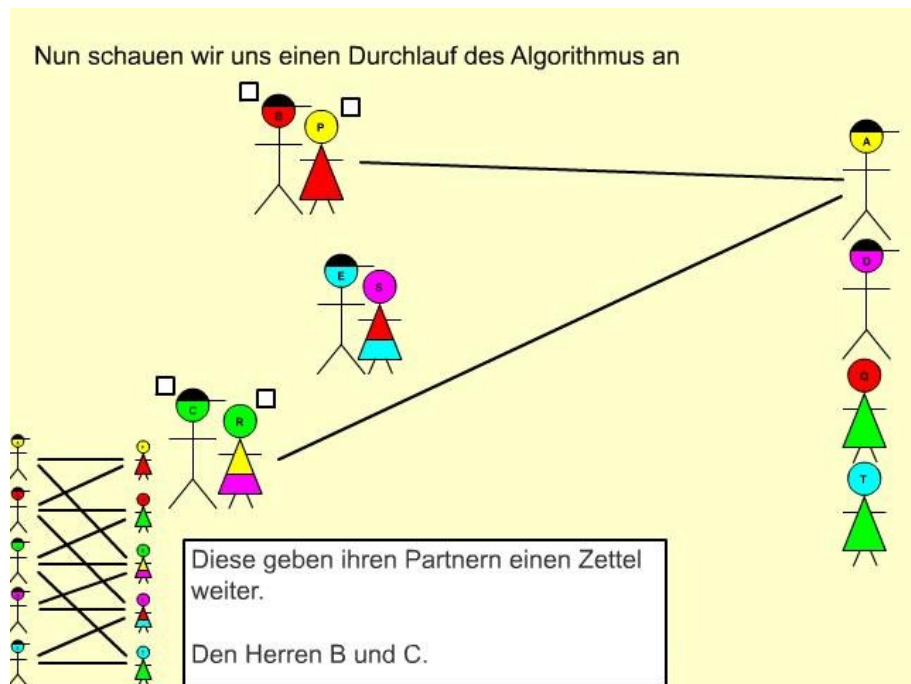
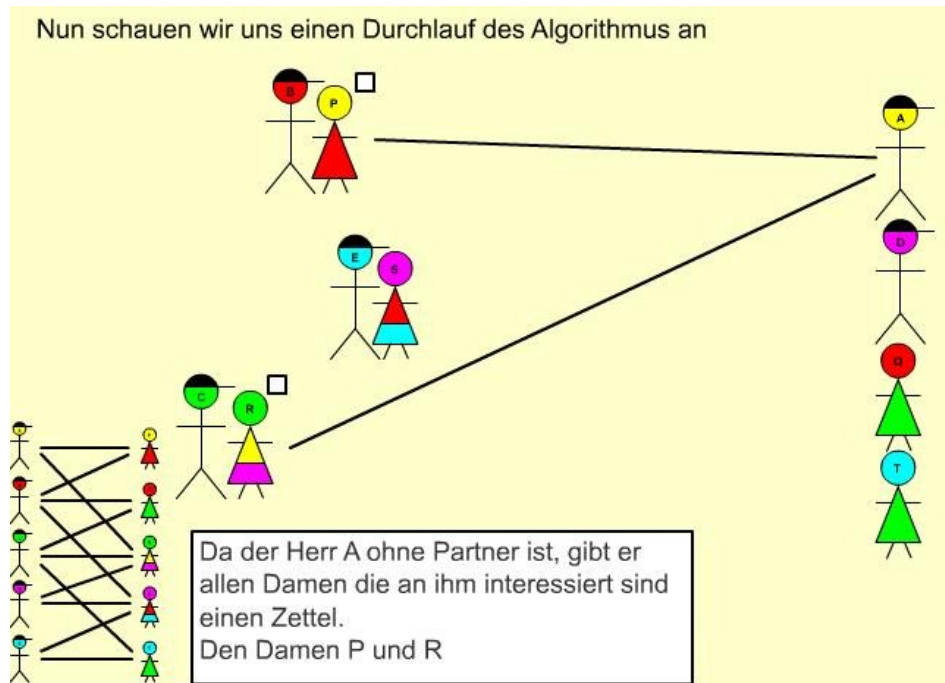
Nun starten wir den “Partnerwechsel“: Eine Person, die noch keinen Partner besitzt, also ein Herr h^* am Rande der Tanzfläche, gibt allen Damen d , mit denen er als (h^*, d) auf der Liste steht, einen Zettel mit seinem Namen. Jede dieser Damen gibt nun einen Zettel mit ihrem Namen an den Herrn, mit dem sie soeben tanzt. Jeder dieser Herren gibt nun einen Zettel mit seinem Namen an jede Dame, die eine Partnerin für ihn sein könnte, allerdings nur, sofern diese mögliche Partnerin noch keinen Zettel hat. Alle diese Damen geben nun erneut einen Zettel an die mit ihnen tanzenden Herren. Hierbei ist es natürlich verboten, dass eine Person zwei Zettel gleichzeitig erhält; vielmehr nimmt jede Dame höchstens einen angebotenen Zettel, wobei es allerdings gleichgültig ist, welchen Zettel sie nimmt, falls ihr mehrere angeboten werden.

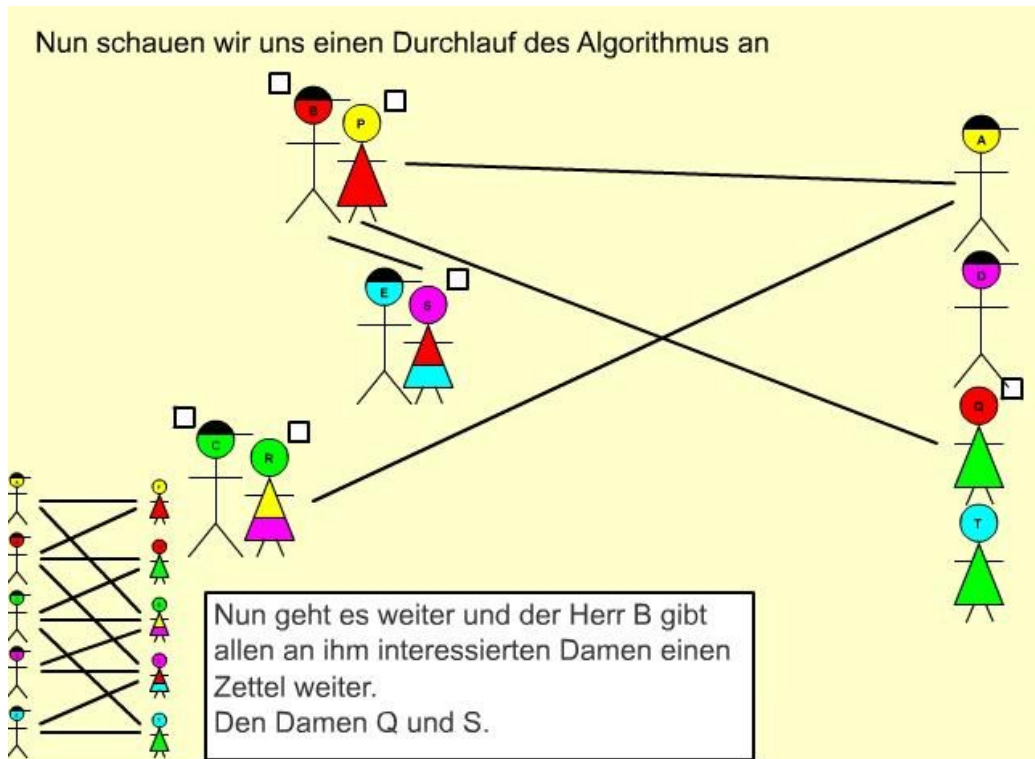
Dieses Verteilen der Zettel geht so lange, bis eine Dame d^* einen Zettel erhält, die noch nicht tanzt. In diesem Augenblick stoppt das Verfahren und diese Dame nimmt als Partner den Herrn, von dem sie den Zettel erhielt, die hierdurch frei gewordene Tanzpartnerin nimmt den Herrn, von dem sie einen Zettel erhielt usw., d.h., es findet eine Kette von Partnerwechseln statt, an deren Ende die Dame d^* und der Herr h^* Partner erhalten haben und insgesamt ein Paar mehr gefunden wurde.

Nun werden alle Zettel weggeworfen und dieser Partnerwechsel wird wiederholt, bis irgendwann das Zettelverteilen dadurch endet, dass keine tanzende Dame mehr einen Zettel erhält und keine außen stehende Dame d^* mehr gefunden werden kann. In diesem Fall kann der Herr h^* keine Partnerin finden und er scheidet endgültig aus der Partnersuche aus. Hat man dies für alle nicht tanzenden Herren durchgeführt, so endet das Verfahren und es wurde eine maximale Teilmenge an Paaren gefunden.

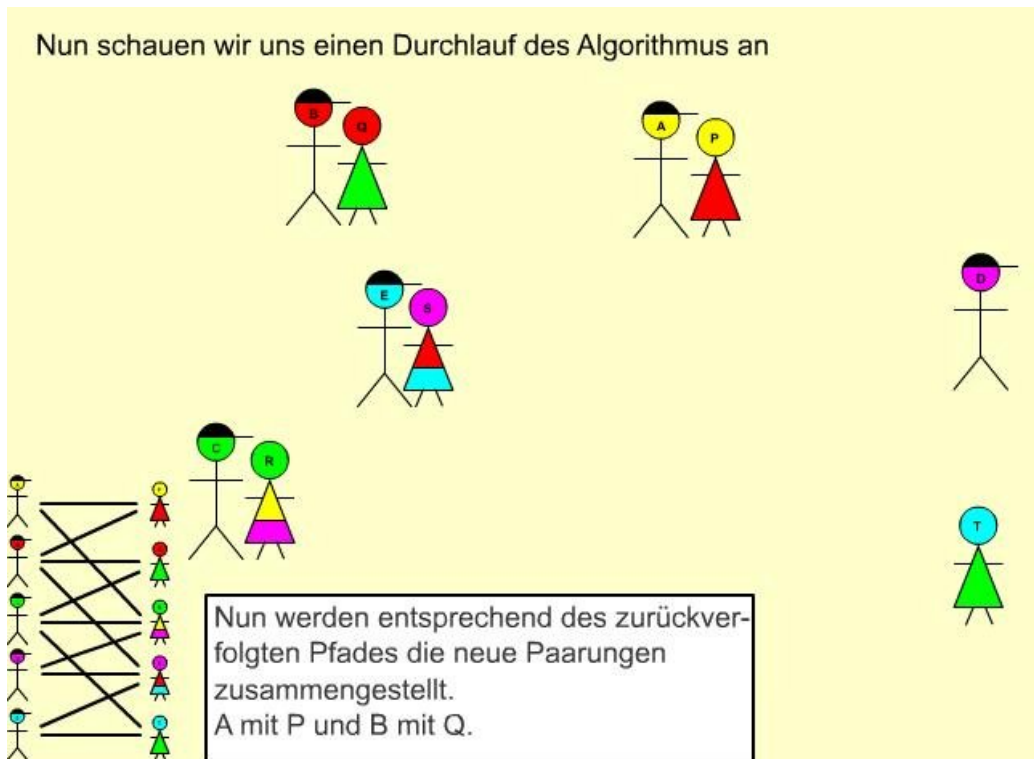
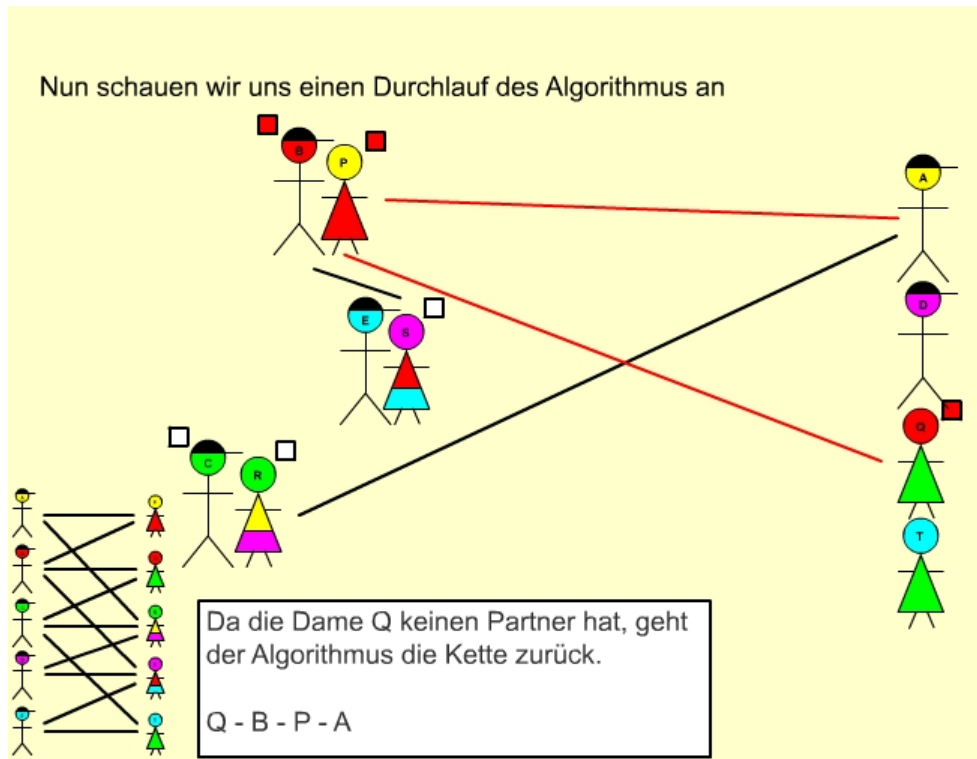
Anschaulich ist der Algorithmus leicht verständlich: Ein Herr, der noch keine Partnerin hat, fragt alle potenziellen Partnerinnen, ob sie ihre derzeitige Partner verlassen könnten. Diese fragen ihre bisherigen Partner, ob diese dann eine neue Partnerin finden würden und diese fragen wieder ihre Partner usw., bis hierbei eine Dame gefragt wird, die noch keinen Partner hat. Dann verfolgt man diese “erfolgreiche“ Fragekette zurück und führt entlang dieser Kette den Partnerwechsel durch.

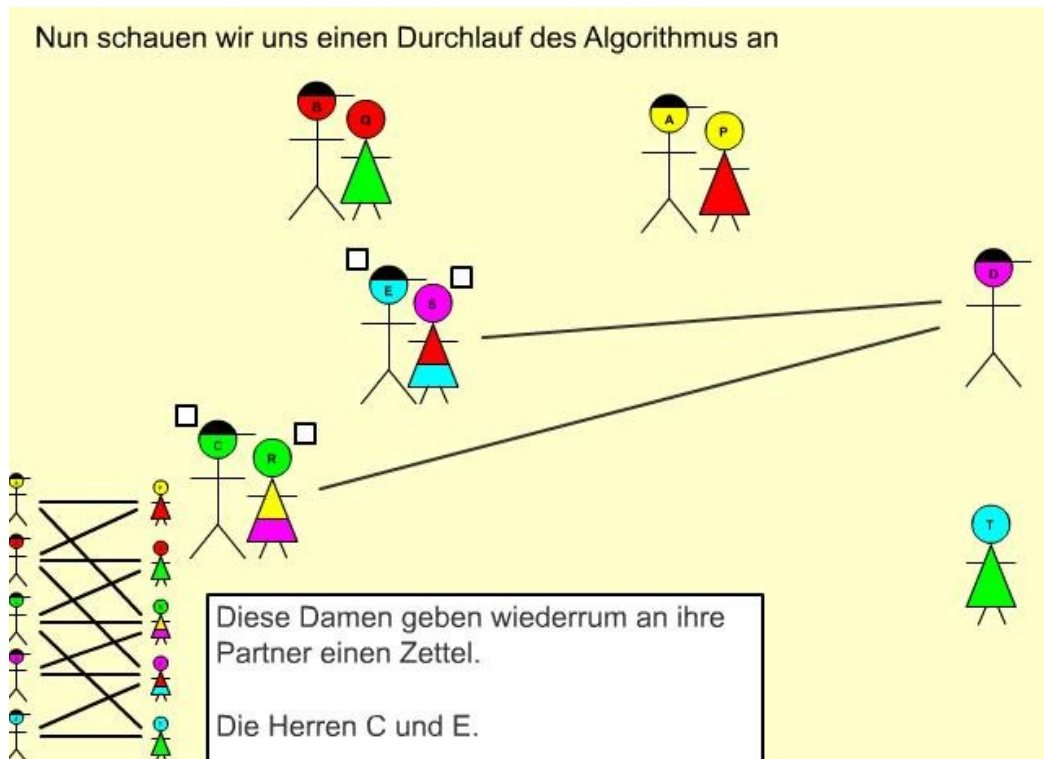
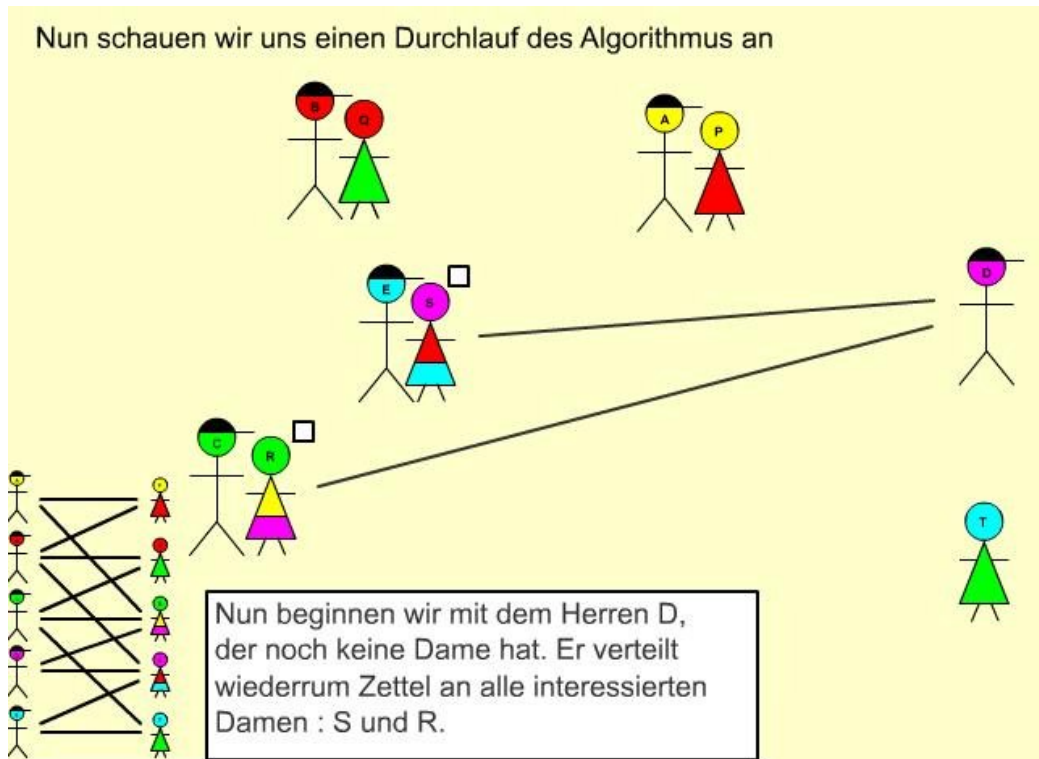
Im genannten Beispiel wurden die Paare (B,P) , (C,R) und (E,S) ausgewählt. Nun beginnt man z.B. mit dem Herrn A. Dieser gibt Zettel an seine potenziellen Partnerinnen P und R. Diese geben danach Zettel an ihre derzeitigen Partner B und C. Nun gibt B je einen Zettel an Q und an S (P hat ja schon einen Zettel), und weil sich hierunter eine Dame, nämlich Q, befindet, die noch nicht tanzt, stoppt das Verfahren und von Q ausgehend wird die Kette der Zettelübergaben rückwärts durchlaufen: Q - B - P - A.

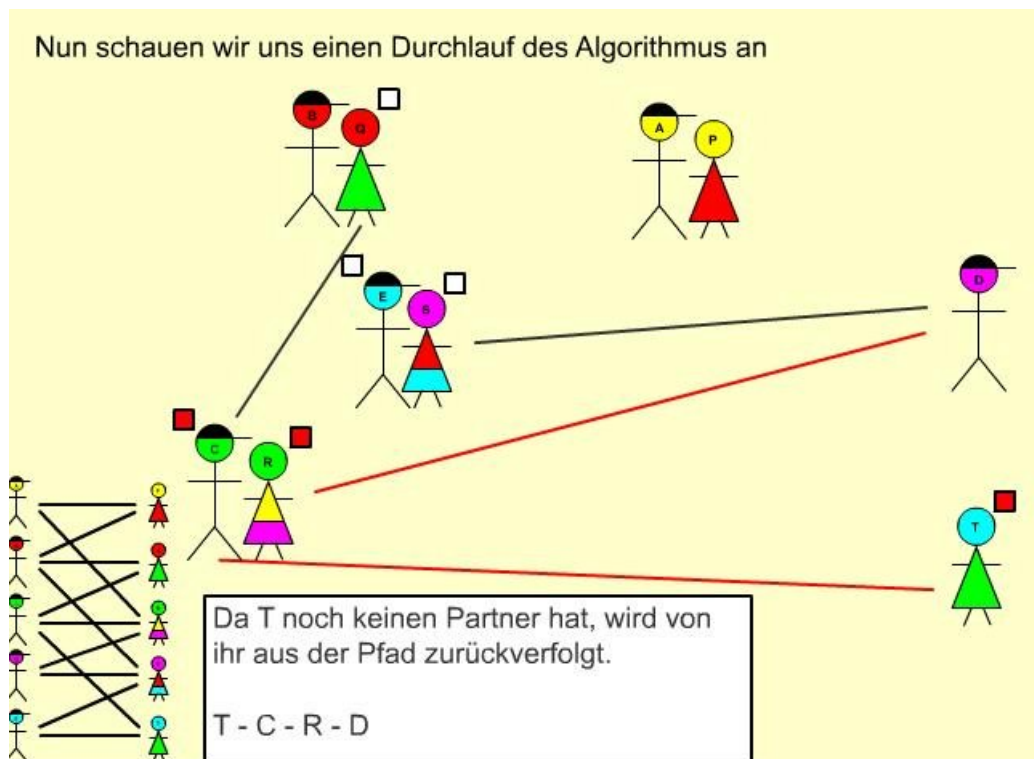
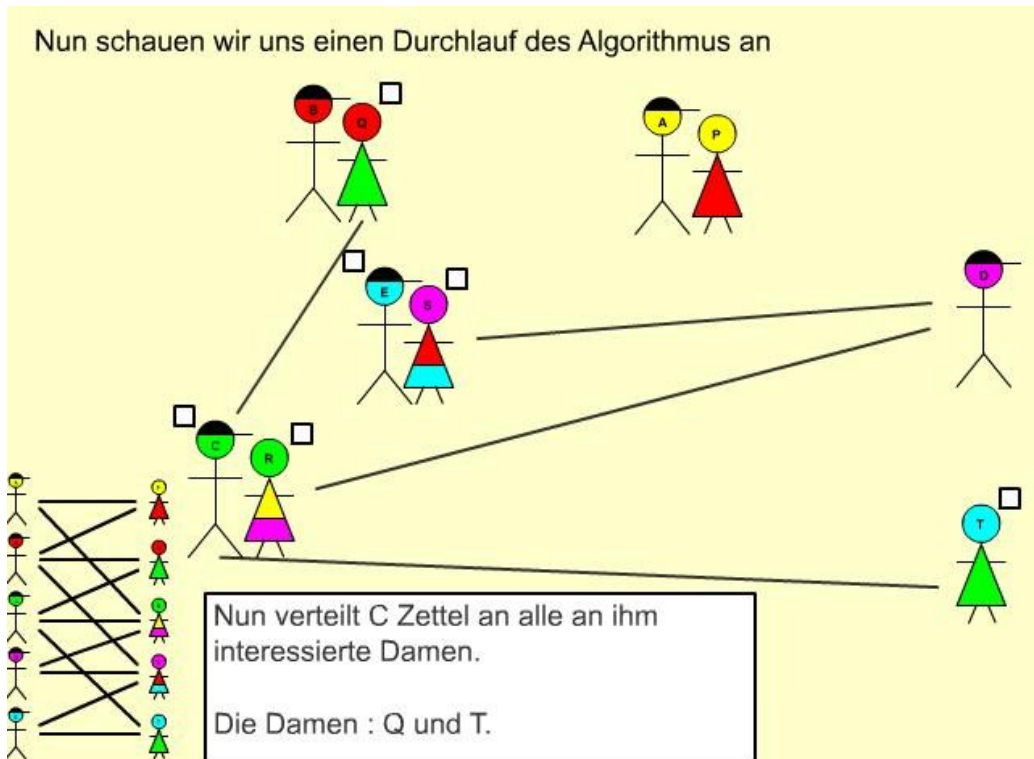


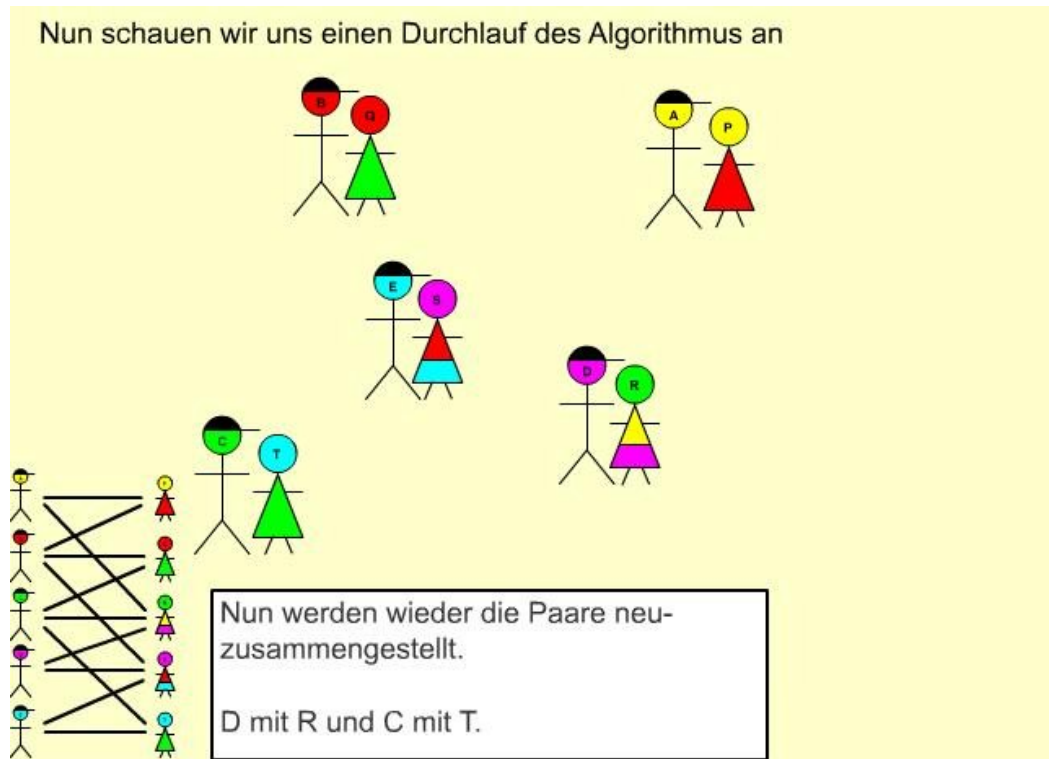


Die Partnerschaften werden nun ausgetauscht, so dass (B, P) wegfällt und dafür (B, Q) und (A, P) entstehen. Auf diese Weise erhält man vier Paare (A, P) , (B, Q) , (C, R) und (E, S) . Nun wird dies Verfahren mit dem Herrn D erneut durchgeführt: D gibt Zettel an R und S , diese geben Zettel an ihre derzeitigen Partner C und E , C gibt einen Zettel an Q und T , und hier endet der Vorgang bereits, da T noch nicht tanzt; wegen der Kette $T - C - R - D$ wird (C, R) durch (C, T) und (D, R) ersetzt, so dass schließlich die Paare (A, P) , (B, Q) , (C, T) , (D, R) , (E, S) entstanden sind.

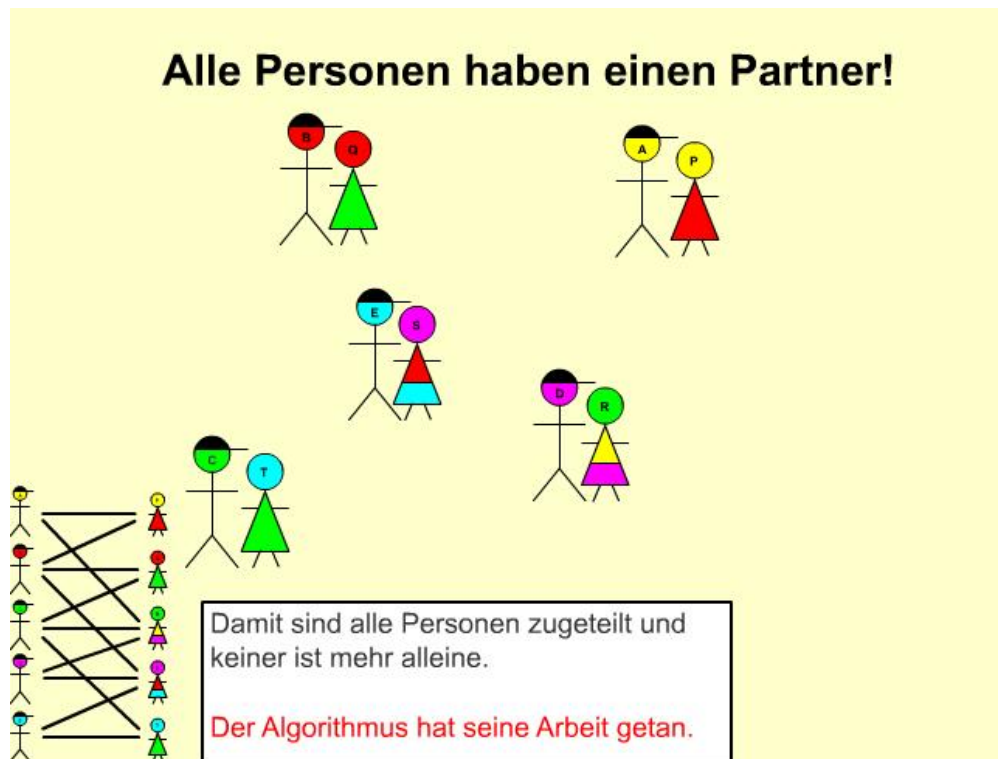








Die Kette muss nicht so kurz sein wie hier angegeben. Wenn wir zum Beispiel von den tanzenden Paaren (B,P) , (C,Q) , (D,R) und (E,S) ausgegangen wären, so würde der verbleibende Herr A je einen Zettel an P und R geben, diese geben Zettel an ihre derzeitigen Partner B und D, B gibt einen Zettel an S (P hat ja schon einen) und D kann keinen Zettel mehr verteilen, weil seine potenziellen Partnerinnen R und S nun schon je einen Zettel haben. Nun gibt S einen Zettel an E und E gibt einen Zettel an T. Da T noch keine Partnerin hat, bilden wir aus der zurückverfolgten Fragekette $T - E - S - B - P - A$ die neuen Paare (E,T) , (B,S) und (A,P) und nehmen diese drei Paare anstelle von (E,S) und (B,P) in unsere Teilmenge auf. So entsteht das Ergebnis (A,P) , (B,S) , (C,Q) , (D,R) und (E,T) . Dies ist ebenfalls eine maximale Lösung. Man sieht, dass es im Allgemeinen mehrere Lösungen eines konkreten Heiratsproblems gibt.



Nun wollen wir den Algorithmus noch ausformulieren. Gegeben sind die Mengen H und D und eine Liste L von r möglichen Partnern der Form $(h_1, d_1), (h_2, d_2), \dots, (h_r, d_r)$. Berechnet wird eine Teilmenge T von maximal vielen Paaren, wobei jedes Element aus H und D höchstens einmal vorkommt.

PARTNERVERMITTLUNG

```

1  T := ∅   { leere Menge }
2  while es noch einen Herrn h* in H gibt, der nicht in T vorkommt do
3    verteile von h* ausgehend schrittweise Zettel wie oben angegeben;
4    if hierbei erhält eine Dame d*, die bisher noch nicht in T vorkommt, einen Zettel then
5      ermittle die Kette d* = d'_n - h'_n - d'_{n-1} - h'_{n-1} - ... - h'_2 - d'_1 - h'_1 = h*;
6      entferne (h'_n, d'_{n-1}), ..., (h'_3, d'_2), (h'_2, d'_1) aus T;
7      füge (h'_n, d'_n), (h'_{n-1}, d'_{n-1}), ..., (h'_2, d'_2), (h'_1, d'_1) zu T hinzu;
8    else entferne h* aus der Menge H end if;
9  endwhile;
10 return T;

```

Der Heiratssatz

Dass das Verfahren korrekt arbeitet, beruht auf dem Heiratssatz des englischen Mathematikers Philip Hall aus dem Jahr 1935. Nach seinem Satz existiert eine Zuordnung aller Herren innerhalb einer Gruppe zu passenden Damen genau dann, wenn folgende Bedingung gilt:

Für jede Teilmenge von Herren gibt es eine mindestens ebenso große Teilmenge von möglichen Partnerinnen.

Die Bedingung bedeutet, dass wenn z. B. 17 Herren betrachtet werden, dann auch mindestens 17 potenzielle Partnerinnen für sie zur Verfügung stehen. Statt 17 dürfen wir auch jede andere Zahl einsetzen. Hierbei geht es zunächst nur um das richtige Verhältnis der Zahlen. Auf die Zuordnung kommt es noch nicht an.

Das Kriterium des Heiratssatzes eignet sich nicht direkt, um auch eine Lösung zu finden. Erstens würde die Überprüfung bei je 50 Damen und Herren mehr als eine Million Jahre dauern, selbst wenn pro Sekunde eine Milliarde Kombinationen von Herren geprüft werden könnten. Zweitens liefert der Satz nur die Information, ob eine Lösung existiert, aber nicht wie sie aussieht. Hierfür benötigen wir den eben beschriebenen Algorithmus.

Traumhochzeit

Man kann mit formalen Methoden beweisen, dass unsere Methode tatsächlich eine maximal mögliche Zahl von Paaren erzeugt. Gilt zu Anfang die Heiratsbedingung, so kann man zeigen, dass kein Herr nach Hause gehen muss und am Ende also alle tanzen und dann eine passende Kandidatin zugeteilt bekommen.

Zeitanalyse

Das Verfahren wird durch die Kettenbildung in Runden aufgeteilt. Um die Analyse zu vereinfachen, nehmen wir an, dass es sowohl n Herren als auch Damen gibt und dass für jeden Herrn k Kandidatinnen möglich sind. Der Parameter $m = n \cdot k$ beschreibt, wie viele Kombinationen von Damen und Herren im Prinzip berücksichtigt werden müssen. Jede Runde kann dann in einer Anzahl von Schritten ausgeführt werden, die zu m proportional ist. Weil nach jeder Runde entweder ein Herr ausscheidet oder ein weiteres Paar hinzukommt, tanzen nach n Runden alle verbleibenden Herren. Als Laufzeitabschätzung erhalten wir also $m \cdot n$ als die Gesamtzahl der Schritte, die während des Algorithmus ausgeführt werden.

Gibt es schnellere Methoden, um eine maximale Menge von Paaren zu berechnen? Ein Indiz für Zeitverschwendung in unserem Algorithmus ist, dass viele Zettel von den Tänzern achtlos weggeworfen werden. Wird die Information dieser Zettel genutzt, können parallel möglichst viele neue Paare gebildet werden. So lässt sich eine Laufzeit proportional zu $m \cdot \sqrt{n}$ erreichen, also spart man den Faktor \sqrt{n} ein. Dieses beschleunigte Verfahren wurde schon 1971 von den amerikanischen Forschern John E. Hopcroft und Richard M. Karp entwickelt, die später jeweils den Turing Award erhalten haben, eine Art Nobelpreis der Informatik.

Autoren:

- Prof. Dr. Volker Claus
<http://www.informatik.uni-stuttgart.de/fmi/fk/mitarbeiter/claus/>
- Prof. Dr. Volker Diekert
<http://www.informatik.uni-stuttgart.de/fmi/ti/mitarbeiter/Diekert/>
- Dr. Holger Petersen
<http://www.informatik.uni-stuttgart.de/fmi/ti/mitarbeiter/Petersen/>

Weiterführende Literatur:

- Dexter C. Kozen: The Design and Analysis of Algorithms. Springer, 1992. (Kapitel 19 und 20)