

Dr. Alexander Kröller

Algorithm Engineering Übung 1 vom 20. 4. 2009

Abgabe der Lösungen am Montag, den 11. 5. 09, vor der Übung im IZ 251.
 Bitte die Blätter vorne deutlich mit eigenem Namen versehen!

Aufgabe 1 („Tobias’ Doodle“): Tobias will mit seinen zwölf Praktikumsteilnehmern regelmässige Termine veranstalten. Daher startet er ein Doodle und lässt es von allen ausfüllen:

	Mo. 08:00	Mo. 09:45	Mo. 11:30	Mo. 13:15	Mo. 15:00	Mo. 16:45	Mo. 18:30	Di. 08:00	Di. 09:45	Di. 11:30	Di. 13:15	Di. 15:00	Di. 16:45	Di. 18:30	Mi. 08:00	Mi. 09:45	Mi. 11:30	Mi. 13:15	Mi. 15:00	Mi. 16:45	Mi. 18:30	Do. 08:00	Do. 09:45	Do. 11:30	Do. 13:15	Do. 15:00	Do. 16:45	Do. 18:30	Fr. 08:00	Fr. 09:45	Fr. 11:30	Fr. 13:15	Fr. 15:00	Fr. 16:45			
OK	OK				OK	OK	OK		OK	OK		OK		OK	OK		OK	OK	OK	OK					OK	OK											
OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK				OK	OK		OK	OK	OK			OK	OK				OK	OK		OK					OK	OK	
			OK	OK	OK									OK	OK							OK	OK														
			OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK					OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	
OK			OK			OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK											OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	OK	OK			OK																						OK	OK									
OK	OK	OK			OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK											OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
OK	OK	OK			OK																						OK	OK									
Tobias	OK	OK	OK	OK	OK			OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK
Count	7	8	7	8	6	10	10	10	4	9	6	6	8	8	8	8	7	9	9	10	11	7	3	3	7	8	8	4	4	5	8	6	5	4			

Tobias beschliesst, vier Termine anzubieten und die Teilnehmer darauf zu verteilen. Natürlich muss zu jedem Termin sowohl Tobias Zeit haben als auch die jeweiligen Teilnehmer. Er hätte gerne eine möglichst gleichmässige Verteilung der Teilnehmerzahlen, d.h., am schönsten wäre es, wenn zu jedem Termin drei Studenten kommen. Wenn das nicht geht, wäre eine maximale Gruppengrösse von vier okay, sonst eben fünf, usw.

- a) Formuliere Tobias’ Problem als IP. Beschreibe die einzelnen Bestandteile des IPs und begründe, warum Dein IP das Problem korrekt modelliert. **(10 P.)**

- b) Löse Dein IP mittels SCIP oder CPLEX. Der oben gezeigte Datensatz liegt als `doodle-real.txt` auf der Vorlesungshomepage. Die Daten sollen mit ZIMPL oder einem selbst geschriebenen Programm in das LP-Format überführt werden.
 Zur Abgabe dieses Aufgabenteils gehören: (1) Konvertierungscode, (2) optimale Lösung, (3) Angabe der Laufzeit der IP-Solvers.
(Zur Herausforderung: SCIP löst mein Modell optimal in 0.04 Sekunden.) **(10 P.)**

- c) Implementiere einen eigenen deterministischen¹ Algorithmus, der versucht, eine möglichst gute Lösung zu konstruieren. Er muss nicht eine optimale Lösung berechnen können. Beschreibe den Algorithmus und die Idee dahinter. Implementiere den Algorithmus und lasse ihn auf `doodle-real.txt` los.

¹d.h. ohne Zufallseinfluss

Zur Abgabe dieses Aufgabenteils gehören: (1) Programmcode, (2) gefundene beste Lösung, (3) Angabe der Laufzeit in mit Teil b) vergleichbarer Umgebung. **(10 P.)**

d) Implementiere einen eigenen randomisierten Algorithmus, der versucht, eine möglichst gute Lösung zu konstruieren. Beispiele solcher Algorithmen:

- Erzeuge massenhaft Zufallslösungen (also vier Termine und zufällige Zuweisung der Teilnehmer), bis sich abzeichnet, dass sich die beste bekannte Lösung nicht mehr verbessert.
- Erzeuge eine Zufallslösung. Evaluiere mögliche Verbesserungsschritte (z.B. Tausch eines Termins, Tausch von Teilnehmern, usw.) und führe einen zufälligen davon durch. Wiederhole, bis sich die beste bekannte Lösung nicht mehr verbessert.

Dein Algorithmus muss nicht eine optimale Lösung berechnen können. Beschreibe den Algorithmus und die Idee dahinter. Implementiere den Algorithmus und lasse ihn auf `doodle-real.txt` los.

Zur Abgabe dieses Aufgabenteils gehören: (1) Programmcode, (2) gefundene beste Lösung, (3) Angabe der Laufzeit in mit Teil b) vergleichbarer Umgebung. **(10 P.)**

e) Evaluiere das Laufzeitverhalten der Algorithmen aus b), c) und d) anhand von Zufallsdaten (`doodle-r1.txt` bis `doodle-r9.txt`). **(10 P.)**

Zum Datenformat der Beispieldaten:

- In der ersten Zeile steht, wie viele Termine Tobias anbietet.
- Danach kommt für Tobias sowie jeden Studenten eine Zeile, in der die Einträge mit einem Tab getrennt sind. Vorne steht „Tobias“ oder eine Studenten-ID, gefolgt von 34 Einträge für die 34 möglichen Zeitslots. Der Eintrag ist 1, wenn die Person zu diesem Zeitpunkt kann, sonst 0.
- In der letzten Zeile steht nur „end“.

($\Sigma=50$ P.)

Aufgabe 2 (Schule): Eine Berliner Realschule steht jedes Jahr vor demselben Problem: Eine grössere Menge Ex-Grundschüler muss auf die Klassen 7a bis 7d verteilt werden (in Berlin ist die Grundschulzeit 6 Jahre). Dazu setzen sich zwei Tage lang die jeweiligen Klassenlehrer an einen runden Tisch und dealen untereinander mit Anmeldungen, bis sie am Schluss eine halbwegs faire Aufteilung haben (und heillos zerstritten und fertig mit den Nerven sind).

Dieses Jahr wollen sie eine externe Aufteilung, die natürlich vollkommen fair sein soll. Zu den Fairnessbedingungen gehören folgende Aspekte:

- (F) Je höher der Mädchenanteil, desto ruhiger die Klasse. Daher sollten alle Klassen möglichst gleich viele Mädchen haben.
- (H) Schüler, die trotz Hauptschulempfehlung auf der Realschule angemeldet wurden, sind oft betreuungsintensiv. Daher sollten alle Klassen möglichst gleich viele dieser Schüler haben.
- (M) Das Verhältnis von Schülern mit Migrationshintergrund zu solchen ohne sollte in allen Klassen etwa gleich sein.
- (N) Der Durchschnitt der Grundschulabgangsnoten sollte in den Klassen möglichst gleich sein.

Natürlich kann nicht jeder Schüler in jede der vier Klassen gesteckt werden. Es gibt einige Rahmenbedingungen:

1. Die Klassenstärken sind vorgegeben: In der 7a und 7c sollen je 26 Schüler sein, in der 7b und 7d jeweils 27.
2. Wahlpflicht französisch wird nur in der 7c und 7d unterrichtet. Schüler, die dieses Wahlpflichtfach gewählt haben, müssen also in einer dieser Klassen.
3. Ähnliches gilt für den christlichen Religionsunterricht (beider Konfessionen). Diese Schüler müssen in die 7a oder 7c. Schüler anderer Glaubensrichtungen können in jede Klasse.

Zur Modellierung dieses Problems: Offensichtlich kann man nicht vier Zielfunktionen (F, H, M, N) haben. Eine sinnvolle Strategie ist hier, drei der vier zu Nebenbedingungen zu machen. Bei (F) etwa „Die Anzahl der Mädchen in jeder Klasse muss zwischen F_{\min} und F_{\max} liegen“. Meiner Ansicht nach ist es sinnvoll, F, H und M zu Nebenbedingungen zu machen und N zur Zielfunktion. Den Lehrern ist das aber eher schnuppe, solange eine nette Klassenaufteilung herauskommt, man kann also von diesem Vorschlag abweichen.

- a) Formuliere das Schulproblem IP. Beschreibe die einzelnen Bestandteile des IPs und begründe, warum Dein IP das Problem korrekt modelliert. **(15 P.)**
- b) Löse Dein IP mittels SCIP oder CPLEX. Der oben gezeigte Datensatz liegt als `schule.txt` auf der Vorlesungshomepage. Die Daten sollen mit ZIMPL oder einem selbst geschriebenen Programm in das LP-Format überführt werden.

Zur Abgabe dieses Aufgabenteils gehören: (1) Konvertierungscode, (2) optimale Lösung, (3) Angabe der Laufzeit der IP-Solvers. **(20 P.)**

- c) Implementiere einen eigenen Algorithmus, der versucht, eine möglichst gute Lösung zu konstruieren. Er muss nicht eine optimale Lösung berechnen können. Beschreibe den Algorithmus und die Idee dahinter. Implementiere den Algorithmus und lasse ihn auf `schule.txt` los.

Zur Abgabe dieses Aufgabenteils gehören: (1) Programmcode, (2) gefundene beste Lösung, (3) Angabe der Laufzeit in mit Teil b) vergleichbarer Umgebung. **(15 P.)**

Zum Datenformat:

- In der ersten Zeile stehen zur Erinnerung Spaltenköpfe.
- Danach folgt für jeden der 106 Studenten eine Zeile, in der mit Tab getrennt folgende Einträge stehen:
 - 1 ID des Schülers.
 - 2 Kann er in die 7a gesetzt werden? (1 oder 0, je nach Reli bzw. Franz)
 - 3 Kann er in die 7b gesetzt werden? (1 oder 0, je nach Reli bzw. Franz)
 - 4 Kann er in die 7c gesetzt werden? (1 oder 0, je nach Reli bzw. Franz)
 - 5 Kann er in die 7d gesetzt werden? (1 oder 0, je nach Reli bzw. Franz)
 - 6 Ist es ein Mädchen? (1 oder 0)
 - 7 Hat er nur eine Hauptschulempfehlung? (1 oder 0)
 - 8 Hat er einen Migrationshintergrund? (1 oder 0)
 - 9 Grundschulabgangsnote als Punktzahl (je höher desto besser)
- In der letzten Zeile steht nur „end“.

($\Sigma=50$ P.)