Data-Driven Decision Making
Assignment 3: Improving Kidney Exchange Programs

João Pedro PEDROSO

DCC - FCUP, May 2025

In this assignment there will be some questions based on the following exercises. In the assignment's class there will be a set of questions in Codex*, with the computers set up as in previous classes. The AMPL book and the classes' slides will be available for consulting.*

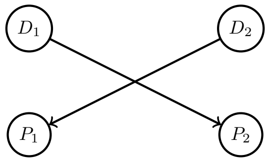
Note 1: *You will be able to use the commercial software AMPL (`https://ampl.com`), independently or via its Python API. A version with a license for this course is available in `https://www.dcc.fc.up.pt/~jpp/AMPL`. A well-known solver for dealing well with integer optimization problems is `gurobi`.*

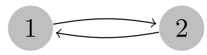
Note 2: *Please bring scratch paper, and do not use any other materials or electronic devices during the class.*

Note 3: *Even though students are encouraged to work in collaboration with a colleague in the preparation of this assignment, each student will have to submit her/his answers separately.*

Renal diseases affect thousands of patients, who, to survive, must incur in dialysis — a costly treatment with many negative implications on their quality of life. As an alternative, patients may enter a waiting list for receiving a kidney from a deceased donor; however, waiting times are typically very long.

For reducing the waiting time, another alternative in some countries is to find a healthy living donor — usually, a relative of a person emotionally connected — who volunteers to cede one of his kidneys. However, in some situations transplantation is not possible due to blood, or tissue-level incompatibility. In these cases, a donor-patient (D_1, P_1) may enter a pool of pairs in the same situation; if in the list there is another pair (D_2, P_2) such that D_2 is compatible with P_1, and D_1 is compatible with P_2, they may do crossed transplants; this can be represented as a cycle in a graph, as shown in the right-hand figure.





Longer cycles may be allowed. In this particular program, compatible pairs may enter the program, in the hope to obtain a better exchange for the patient.

Compatibilities among blood groups are as follows.

	Patient			
Donor	O	A	B	AB
O	✓	✓	✓	✓
A	✗	✓	✗	✓
B	✗	✗	✓	✓
AB	✗	✗	✗	✓

One of the difficulties that doctors face when choosing a matching is assessing the longevity patients will have after receiving a transplant. In this assignment you are asked to help by providing a model that, based on past data, computes for each compatible donor-patient link the expected survival time for the patient if this transplant was made. It has been decided to use a decision tree with a maximum depth of 4. Assume that transplantation is accepted, for patients that are blood-type compatible with the donor, only if 5 or more years of remaining life are expected.

The data currently available concerning past transplants is the following: `historical.csv`. Data contains sex, blood type, age, weight, and height of the donor, the same data for the patient, followed by a series of blood qualities characterizing the patient's condition. The last column is the survival time after transplant. (These data were generated with a method loosely inspired on lkdpi.)

Fields are the following:

donor age	
donor sex	M/F
donor egfr	(see description in lkdpi)
donor sbp	(see description in lkdpi)
donor bmi	body mass index
donor smoker	donor is a smoker (1 if true)
donor relat	donor is a relative of patient (1 if true)
donor abo	ABO blood type for donor
donor weight	
donor hlaB	human leukocyte antigens B (see description in lkdpi)
donor hlaDR	human leukocyte antigens DR
patient age	
patient sex	
patient smoker	
patient abo	
patient weight	
patient hlaB	
patient hlaDR	
survival time	observed survival time for patient after transplant

(As for HLA: having different antigens on donor and patient is not considered good for long term survival.)

There is also data concerning the donor-patient pairs in the current pool in `pool.csv`. Fields are the same as in historical data, without survival time, preceded by an ID of each pair. Note that in the pool some pairs are compatible.

Your task is to find the best donor-patient matching in this pool, i.e., the matching that maximizes the total expected survival time for patients receiving a transplant, subject to the previous conditions. Compute the solution with maximum total survival time.

1. Case 1: no special constraints concerning exchanges involving a compatible pair.
2. Case 2: considering *rationality constraints:* a compatible pair will accept to participate in a crossed transplant only if the expected survival time is better than the internal transplant (*i.e.*, the patient receiving the kidney from their donor).