

MITOCW | MIT15_071S17_Session_9.3.03_300k

Let us demonstrate the fundamentals of the approach that is called integer optimization that is used in the eHarmony matching algorithm.

Suppose we have three men and three women.

Their compatibility scores range between 1 and 5 for all pairs, and they are shown in the figure.

For example, the first man and the first woman have a compatibility score of 1.

The first man and the second woman have a compatibility score of 3, and so forth.

So how should we match pairs together to maximize compatibility?

The red line indicates a match.

In this case man one is matched to the third woman with a compatibility score of 5.

The second man to the first woman with a compatibility score of 4.

And the third man to the second woman with a compatibility score of 5.

And a total compatibility score of 14.

Let us now formulate this matching problem in the language of integer optimization.

We denote the data for the matching problem as w_{ij} .

These numbers represent the compatibility score between user i and user j .

In the example shown in the figure, w_{13} is 5.

We first define decision variables x_{ij} .

These are binary variables taking value 1 if we match user i and user j together, and value 0 otherwise.

We next define the objective function, which is to maximize compatibility between matches.

In this case, we maximize the sum of $(w_{11}x_{11}) + (w_{12}x_{12})$ and so forth, up to $(w_{33}x_{33})$.

We next define the constraints for the decision variables x_{ij} .

Each man-- say, man one --should match to exactly one woman.

And we represent this by the mathematical constraint as follows-- $x_{11} + x_{12} + x_{13} = 1$.

Similarly, each woman-- say, woman one --should match to exactly one man.

And we represent this constraint mathematically as $x_{11} + x_{21} + x_{31} = 1$.

The full optimization problem is indicated next.

First we have the objective function we indicated earlier.

Next we have the set of constraints that indicate that every man should match with exactly one woman.

Next we have the constraints that every woman should match with one man.

And finally that the variables are binary.

Let us now extend the methodology to other areas.

Suppose, for example, we want to show each woman her top two matches.

This is represented by the constraint $x_{11} + x_{21} + x_{31} = 2$.