Lecture Note 15: Fair Allocation with Combinatorial Utilities

Christian Kroer*

April 5, 2022

1 Introduction

Recall that for the setting of indivisible goods, a market equilibrium is not guaranteed to exist. Moreover, envy-free allocations are also not guaranteed to exist. In this lecture note we will see how to recover existence by considering an appropriate notion of approximate market equilibria. We will use this to design a fair method for allocating course seats to students.

Specifically, we will look at a generalization of the *competitive equilibrium from equal incomes* (CEEI) allocation mechanism. Since a market equilibrium is not guaranteed to exist for equal budgets, we will instead look at *approximate CEEI* (A-CEEI). In A-CEEI the idea is to relax two parts of CEEI: (1) we give agents approximately equal, rather than exactly equal, budgets, and (2) we only clear the market approximately.

Let's see how this works with an example. Consider an example where two agents are trying to divide four goods: two diamonds (one large (LD), one small (SD)), and two rocks (one pretty (PR), one ugly (UR)). Say the agents both have utilities such that they can take at most two items, and they prefer bundles in the order

$$(LD, SD) > (LD, PR) > (LD, UR) > (LD) > (SD, PR) > (SD, UR) > (SD) > (PR, UR) > (PR) > (UR).$$

Clearly if budgets are equal we cannot hope to price these items in a way that clears the market, since both agents will always want the bundle with the large diamond if they can afford it. But if we instead give agent 1 a budget of 1.2 and agent 2 a budget of 1, then we can set the prices as follows:

LD	SD	PR	UR
1.10	0.8	0.2	0.1

Now agent 1 wishes to buy (LD, UR) for a total price of 1.2, and agent 2 wishes to buy (SD, PR) for a total price of 1. As long as we decide the budget perturbations in a randomized way this is in some sense fair in expectation, and furthermore we might hope that the budget perturbations are small enough that for instances with more than four items, things look even fairer. Note that the allocation we found satisfies both EF1 and the MMS guarantee. The example also achieves Pareto optimality, but we will in general only guarantee approximate Pareto optimality for A-CEEI for more general valuations.

*Department of Industrial Engineering and Operations Research, Columbia University. Email: christian.kroer@columbia.edu.

2 Approximate CEEI

We will describe the problem in the context of matching students to seats in courses. This setup is used in the *Course Match* software, which is used for matching students at Wharton and several other schools. There is a set of m courses, and each course j has some capacity s_j. There is a set of n students. Each student has a set $\Psi_i \subseteq 2^m$ of feasible subsets of courses that they may be allocated, with each bundle containing at most $k \leq m$ courses (note that this assumes that each student can only consume one unit of a good, even if $s_j > 1$; this is of course reasonable in course allocation, but not for all applications). The set Ψ_i encodes both scheduling constraints such as courses meeting at the same time, as well as constraints specific to the student such as whether they satisfy the prerequisites. The preferences of student i are assumed to be given as a complete and transitive ordinal preference ordering $\succcurlyeq_i$ over Ψ_i. Completeness simply means that for all schedules $x, x' \in \Psi_i$, $x \succcurlyeq_i x'$, $x' \succcurlyeq_i x$, or both. Transitivity means that if $x \succcurlyeq_i x'$ and $x' \succcurlyeq_i x''$ then $x \succcurlyeq_i x''$.

Given a set of prices p for each course, a vector x_i^* is in the demand set for student i if

$$x_i^* \in \operatorname{argmax}_{\succcurlyeq_i} \{x_i \in \Psi_i : \langle x_i, p \rangle \leq B_i\}.$$

In the actual Course Match implementation, $\succcurlyeq_i$ is represented numerically by an utility function for each student, but the A-CEEI theory works for the more general case of ordinal preferences.

Since we have existence issues (these arise both from indivisibility as seen earlier, but also from the very general preference orderings allowed), we resort to an approximation to CEEI:

Definition 1. *An allocation x, prices p, and budgets B constitute an (α, β)-CEEI if:*

1. $x_i \in \operatorname{argmax}_{\succcurlyeq_i} \{x' \in \Psi_i : \langle p, x' \rangle \leq B_i\}$ *for all* i
2. $\|z\|_2 \leq \alpha$, *where* $z \in \mathbb{R}_+^m$ *is defined as* $z_j = \sum_i x_{ij} - s_j$ *if* $p_j > 0$, *and* $z_j = \max(\sum_i x_{ij} - s_j, 0)$ *if* $p_j = 0$
3. $B_i \in [1, 1 + \beta]$ *for all* i

The first condition in (α, β)-CEEI simply says that each student i buys an item in their demand set. The second condition says that supply constraints are approximately satisfied. The third constraint says that all budgets are almost the same, up to a difference of β.

The main theorem regarding (α, β)-CEEI is that they are guaranteed to exist:

Theorem 1. *Let $\sigma = \min(2k, m)$. For any $\beta > 0$, there exists a $(\sqrt{\sigma m}/2, \beta)$-CEEI. Moreover, given budgets $B \in [1, 1 + \beta]^n$ and any $\epsilon > 0$, there exists a $(\sqrt{\sigma m}/2, \beta)$-CEEI using budgets B^* such that $\|B^* - B\|_\infty \leq \epsilon$.*

One major concern with this result is that we are not quite guaranteed a feasible solution. In general the allocation may oversubscribe some courses, though the oversubsciption vector z has bounded ℓ_2 norm. In practice, the bound is relatively modest: First, the bound $\sqrt{\sigma m}/2$ does not grow with the number of agents or number of course seats. Second, in practice students take at most a modest number of courses per semester among a reasonably-small number of courses offered (an example given in the literature is that students take $k = 5$ courses out of 50 courses total at Harvard's MBA program), thus yielding a bound of roughly 11. Technically a single course could be oversubscribed by 11 students, but in practice we expect this to be smoothed out reasonably across many courses.

The proof of the existence theorem is rather involved and relies on smoothing out the market in order to invoke fixed-point theorems. Here we give some intuition for the role that each approximation plays.

As in other discontinuous settings, the main difficulty for existence without approximation is the discontinuity of student demands with respect to price. However, in the course match setting, $\sqrt{\sigma}$ is an upper bound on the discontinuity of the demand of any single agent. To see this, note that a demand x_i has at most k entries set to 1, and so a student can at most drop all courses from x_i and switch to k new courses under their new demand x_i'. At the same time, there's only m courses total, so the change is bounded by $\min(2k, m)$, and thus $\|x_i - x_i'\|_2 \leq \sqrt{\sigma}$.

The second discontinuity issue is to avoid large discontinuous aggregate changes in demand across the students. When budgets are the same, as in standard CEEI, the demand discontinuity across students may occur at the same point in the space of prices. Thus, if this happens, aggregate discontinuity may be on the order of $n\sigma$. With distinct budgets, it becomes possible to change a single student's demand without changing those of other students. For each bundle x, we may think of the hyperplane $H(i, x) = \{p : \langle p, x \rangle \leq B_i\}$ which denotes the boundary between two halfspaces in the price space: those where student i can afford x, and those where i cannot afford x. By having each budget distinct, one can show that in a generic sense, at most m hyperplanes can intersect at any particular point in price space. This implies that aggregate demand changes by at most σm.

The remainder of the proof is concerned with smoothing out the aggregate demands so that a fixed-point existence theorem can be applied to show existence.

2.1 Fairness and Optimality Properties of A-CEEI

Since we are only approximately clearing the market, we do not get Pareto optimality. However, it is possible to show that if we construct a modified market where $\tilde{s}_j = s_j - z_j$, then we have Pareto optimality in that market. Thus, any Pareto-improving allocation must utilize unused supply, which can potentially be used to bound the inefficiency once more structure is imposed on utilities.

Crucially, (α, β)-CEEI does guarantee some fairness properties. If we select $\beta \leq \frac{1}{k-1}$, then EF1 is guaranteed in any (α, β)-CEEI. Furthermore, there exists β small enough such that each student is also guaranteed to receive their $(n+1)$-MMS share, which is their utility if they were forced to partition the items into $n+1$ bundles and take the worst one.

2.2 Practical Course Match Concerns

In Course Match, the representation of $\succcurlyeq_i$ is as follows: the set of feasible schedules Ψ_i is taken as given. Then, student i ranks each course on a scale from $0 - 100$, and is additionally allowed to specify pairwise penalties or bonuses in $-200, 200$ for being assigned a given pair of courses.

2.3 Computing A-CEEI

In general computing an A-CEEIis PPAD complete. This is the same class of problem that general-sum Nash equilibrium falls in. It is conjectured to require exponential time in the worst case, and thus we cannot hope to have nice scalable algorithms like we had for the divisible case.

In practice, A-CEEI is computed using local search. A *tabu search* is used on the space of prices. This works as follows:

1. A price vector is generated randomly
2. A set of "neighbors" are generated using two different generation approaches:
    - "Price gradient:" all the demands under the current prices are added up, and the excess demand vector is treated as a gradient. Then, 20 different stepsizes are tried along the price gradient

- A single item has its price changed, and all other prices are kept the same. The new price on the chosen item is set high enough to stop it from being oversubscribed, or low enough to stop being underscribed. A neighbor is generated for each over or undersubscribed item

3. The best neighbor (among the ones generating a previously-unseen allocation) is selected as the next price vector, and the procedure repeats from step 2 (unless the last 5 iterations yielded no improving prices, in which case the local search stops)

4. Finally, step 1 is repeated with a new random price vector. This repeats until a time limit is reached

In practice this procedure generates an A-CEEI solution with significantly better α and β values than the theory predicts, within roughly two days of computation. In the process, about 4.25 billion MIPs are solved. After an A-CEEI has been generated, additional heuristics are implemented in order to force the solution to not have oversubscription.

3 Historical Notes

A-CEEI was introduced by Budish [1], and an implementation of A-CEEI used at Wharton was given by Budish et al. [2]. The proof of PPAD completeness was by Othman et al. [3].

References

[1] Eric Budish. The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, 119(6):1061–1103, 2011.

[2] Eric Budish, Gérard P Cachon, Judd B Kessler, and Abraham Othman. Course match: A large-scale implementation of approximate competitive equilibrium from equal incomes for combinatorial allocation. *Operations Research*, 65(2):314–336, 2016.

[3] Abraham Othman, Christos Papadimitriou, and Aviad Rubinstein. The complexity of fairness through equilibrium. *ACM Transactions on Economics and Computation (TEAC)*, 4(4):1–19, 2016.