STRUCTURED DISCRETE FAIR DIVISION ALGORITHM FOR ALLOCATING SUBTASKS WITHIN STUDENT PROJECTS

Conference Paper · July 2020

DOI: 10.21125/edulearn.2020.0597

2 authors:

Ana Madevska Bogdanova
Ss. Cyril and Methodius University in Skopje

Monika Simjanoska
Ss. Cyril and Methodius University in Skopje

STRUCTURED DISCRETE FAIR DIVISION ALGORITHM FOR ALLOCATING SUBTASKS WITHIN STUDENT PROJECTS

Ana Madevska Bogdanova, Monika Simjanoska

Ss. Cyril and Methodious University, Faculty of Computer Science and Engineering (MACEDONIA)

Abstract

We are investigating the most appropriate allocation of subtasks within students projects which are compulsory part of the Intelligent Systems course taught at the Faculty of Computer Science and Engineering in Skopje, N. Macedonia. The past experience has shown that the students tend to choose the subtasks they find to be the easiest. To avoid the "easiest route" approach, we propose a fair division strategy for solving this problem. The main goal is to guarantee the success and high quality of the student projects. To achieve the aims, our methodology is focused on objective determination of the students' preferences for each project subtask and map them in the most appropriate one. The results have shown that our approach ensures efficiency and envy-freeness, whereas the equitability is partially proved due to the subjective student's evaluation of the personal preferences as a result of the biasness to the teachers valuation of the subtasks.

Keywords: Subtask division, Fair division, Grading.

1 INTRODUCTION

The fair division and collective well-fare have been subjects of discussion since the beginning of the civil society [1]. The first notes of formalizing the definition of distributive fairness was given by Aristotle - two persons having the same characteristics regarding a given allocation problem, should be treated equally [2]. The origins of fair division have been discussed by Brams and Taylor [3] presenting the story of King Solomon's solution to discover the real mother achieve her goal of getting the baby. This, however, is the simplest case with easily predictable outcome. From then on, the strategies have become more complicated, involving more than two agents who have different affinities over divisible/indivisible, homogeneous/heterogeneous kinds of goods.

The first mathematical algorithm to fairly cut a cake among more than two agents was proposed by Steinhaus, Knaster and Banach in 1940s [4]. This strategy is referred to as cake-cutting and is only about continuous resources distribution to individuals and never considers discrete goods [5].

This research is about discrete subtasks fair division within student projects - a problem that cannot be solved by a standard cake-cutting approach, but nevertheless must fulfill the general features of fair division as envy-freeness, equity and Pareto optimality [6]. Opposite of the general cake-cutting strategy where the agents involved aim to satisfy the appetite by maximizing the portion assigned [7], in our case we deal with agents (students) that have the affinity either to minimize the work or to delegate it to the other members of the team.

The project tasks can be divided infinite number of homogeneous subtasks. However, the problem gets more complicated by the fact that the teams are comprised of students with different course performance background. Allowing the students to choose a subtask by their own preferences is very likely to end up with insincere choice that is not consistent with their true preferences (skills) [8].

Hereupon, we want to explore the questions - what are the true preferences and how can we guarantee fairly subtasks assignment and achieve envy-free and efficient team work?

We created an original methodology that analyze students activities during the course teaching and predicts the most suitable portion of cake - the subtask that need to be assigned. The appropriate assignment of subtask is crucial for achieving equality according their skills among the team members and hence, envy-freeness.

In order to have a fully automated procedure, we introduce a human factor for rating the discrete subtasks at the beginning. The subtasks are assigned to the agents according to one-to-one relationship. The assignment is obtained by a valuation function that maps the student's performance

ISBN: 978-84-09-17979-4

into the most suitable subtask. The appropriate assignment of subtasks is crucial for sustaining the successful project outcomes and the didactical quality of the course.

The rest of the paper is organized as follows. The problem for fair division of project subtasks and details of the developed methodology have been presented in Section 2. An empirical evaluation of the proposed strategy is shown in Section 3 where the methodology is tested. In Section 4 we conclude the contribution of our work and provide some interesting directions for future work.

2 METHODOLOGY

2.1 Related work

The latest results on fair division research show that it is especially popular in games and economy. [9] is a book focused on economics and computation provides an insight into the game theory, social choice and fair division.

Considering our interests in discrete allocation problem, we present the work of Aziz et al. [10] who work in the field but instead of deriving the agents preferences, they discuss a case where the agents express their ordinal preferences over objects. The paper is focused on proportionality and envy-freeness as important fair division features. Ranking is commonly used approach in procedures for dividing indivisible (discrete) items. Brams et al. [11] also propose ranking the items from best to worse, but each player must not know the other player’s ranking. Later, Brams et al. [12] propose modified strategy on the same problem that requires each person to rank the items and then finds the allocation such that both of the persons receive the same number of items.

Bouveret and Lemaitre [13] investigate five different fairness criteria and the connection between them in order to characterize how conflicting the agent’s preferences are in a case of allocating indivisible goods. The criteria create an ordered scale that can be used for finding satisfactory fair allocations and measuring the possibility to find some.

Testing fair division algorithms is not very common, say Dupius-Roy and Gosselin in their paper [14] where they provide an empirical evaluation of their fair-division algorithms by performing experiments of satisfaction of two pairs of players who divided 10 indivisible goods between themselves. The results showed that the divisions found by the genetic algorithm they used were rated as more satisfactory than the divisions derived from other six fair-division algorithms.

An interesting paper that demonstrates the applicability of a proposed fair division framework is presented by Porras-Alvarado et al. [15]. It is about performance-based resource allocation method that uses utility functions to al- locate resources in such a manner that participants are convinced to have received a fair share. The model is tested on real data from the Texas Department of Transportation.

An example where fair division is based on Machine Learning approach is presented in the work of Zemel et al. [16]. They present a learning algorithm for fair classification that achieves both group and individual fairness. Their model represents each individual as a data point in a given space and maps each of them to a probability distribution in a new representation space. Hereafter they make new classifications based on the new representations. In their work they point out some recent researchers attempts to achieve group or individual fairness by using Machine Learning techniques [17,18,19].

2.2 Problem definition

The allocation of the subtasks within the projects is mostly a black box for the teacher. Following a distributed approach for allocation, the students were sharing the subtasks by their own beliefs of the personal skills and motivation.

This strategy has shown inconsistency in terms of the project's quality and does not guarantee that it will be successfully finished. The experience and the results from the past year proved that making an inappropriate allocation of subtask may easily cause the student abandoning the project, or even cheating. The first one leads to incompleteness, compelling the other team members do the job, and thus may lead to unsuccessful realization. The cheating might be even worse, meaning that the student is refusing to finish the subtask within the team without teacher's knowledge. Usually, the other team members finish the incomplete subtask, which directly affects the quality of the project.

In either situation, the distributed way of allocation and the personal evaluation of the skills, threaten the project's quality and success.

2.3 Setting the scene

Bouveret and Lang in their paper [1] indicate the following parameters to be very useful when setting the scene for developing an appropriate fair-division strategy:

- What is the nature of the resources to be allocated?
- What is the nature of the preferences of the agents?
- What is the nature of the permitted allocations?
- How to evaluate the quality of an allocation?
- What is the nature of the process that leads to the allocation?

Discussing the nature of the resources is the core of the research problem at hand. We define a course project to be a divisible resource that results in five unequal units - subtasks that can no further be divided. Thus, we deal with a problem of indivisible (discrete) allocation of unequal resources (homogeneous subtasks), meaning that each subtask requires different effort (skills) to be solved.

The nature of the student's preferences over the unequal subtasks is the most challenging part of this research. Under real circumstances, the students have tendency to choose the easiest subtasks. If we follow the cut and choose approach, that would be a subject of a personal evaluation which is far from fair regarding the rest of the team.

Given the fact that each student must be assigned exactly one subtask, we propose original strategy for a fair allocation of subtasks. The first part of the strategy is the students profiling taking into consideration the student's activities during the semester. The profiling scores will then be used as an input in valuation functions that predicts the student's preferences and is later used to find the most fair allocation of subtasks. Thus, the preferences are of numerical nature and are determined by the methodology instead.

The next question is how to evaluate whether the obtained allocation of the subtasks within a project is the best? In [6], a perfect division is considered the one that satisfies the following three properties:

1 Efficiency (Pareto-optimality): There is no other division that is more fair for at least one student without making it worse for at least one student.

2 Envy-freeness: Each student feels that his/her allocated subtask is at least as good as any other share.

3 Equitability: Each student's personal preference for a subtask is the same as the allocated.

Therefore, a successful allocation would be the one that maximizes the previous three properties.

Usually, the nature of the process that leads to appropriate allocation could be either centralized or distributed [20]. A distributed allocation is not convenient in the case we elaborate, since the students (agents) are not reliable to negotiate over the subtask's distribution. Knowing the size of each subtask in prior, one can easily rank the subtasks. This leads to a centralized nature of allocation, providing the teacher with responsibility to decide on each subtask value and rank them accordingly.

2.4 Material and methods

2.4.1 The materials

Each course project is assigned to a team of students (agents). Since the projects consist of five indivisible subtasks, each team must be comprised of exactly five students. The projects' subtasks are numerically ranked by the teacher considering their contribution for the overall project success. Therefore, the subtasks are ordered in the following manner:

1 Management (30%): This subtask is on top of the stack and is rated as most important for successful project. The student who is assigned to be the project manager is supposed to be responsible for organizing the team members, be aware of the deadlines, take care of the results and write a detailed documentation. Given the responsibilities, this student is expected to have high emotional intelligence and advanced writing skills.

2 Data Analysis (25%): This subtask is valued as the second most important since it involves the results analysis and techniques evaluation. It is the key for deriving reliable conclusions.

3 Data Processing (20%): This is the third most important subtask and a student's responsibilities are to choose the right techniques to process the given data. It is usually about reducing the number of features of a given data set and applying some classification or regression techniques.

4 Related Work (15%): This subtask requires high reading skills. The student must be able to recognize how is the other researchers' work related to the project and how the results can be compared once the project is finished.

5 Data Preprocessing (10%): This subtask requires less technical skills and is only about data retrieval and applying some normalization methods if needed.

The success of each subtask highly depends on the others subtasks success. Their natural order would be: Management, Related Work, Data Preprocessing, Data Processing and Data Analysis.

2.4.2 The methods

The student's preferences can be obtained by an approach that we refer to as students Profiling. Recently, we worked on an intelligent modelling of the students' knowledge collected from the e-Learning and e-Assessment processes of a particular course. The early attempt was to adapt the student's e-Learning individual Profile and provide approximation of the reliability of the student's e-Assessment results [21]. Later we proposed a generative modelling of the probability distributions of the students' Profiles that have passed and of those that have failed the course. The probability distributions were applied in the Bayes theorem to classify the students into a pass or fail category [22]. Eventually, the algorithm obtained its advanced form of an intelligent virtual teacher, able to predict the student's most likely final grade at the end of the semester [23].

Using this approach as an inspiration, we propose the following procedure.

2.4.3 The profiling

In the context of the research, the students are depicted in terms of vectors representing the success of their activities during the course up to the moment of assigning the projects. We consider the results from the first partial exam (PE), classes attendance (CA), laboratory exercises (LE), essay on a course topic (ES) and quantum of emotional intelligence (EQ). All but the EQ attributes are objective scores with no human factor. The EQ is the only subjective value assigned by the teacher's personal opinion and describes the student ability to present, communicate and organize.

Hence, let

$S = \{\hat{s}_1, ..., \hat{s}_n\}$ be the set of students where each student is defined as
$s_i = (PE, CA, LE, ES, EQ)$, for $i = 1, ..., n,$ n is the total number of course participants.

Having the students' activity profiles we can determine their responsibility and technical skills represented as their preferences.

2.4.4 Modeling the preferences

In fair division, preferences are usually assumed to be valuation functions [24]. In this case we propose two valuation functions $v_1(s_i)$ and $v_2(s_i)$ to model the i^{th} student responsibility and technical preferences, correspondingly.

Each valuation function values the activity profile attributes differently. The responsibility is more represented through the student's devotion to the course and the ability to communicate and collaborate. Therefore, we consider the fol- lowing order by importance for this feature: CA, LE, EQ, PE, ES. Considering this order, the coefficients in Equation 1 are k_{11} = 0.15, k_{12} = 0.30, k_{13} = 0.25, k_{14} = 0.10 and k_{15} = 0.20.

$$v_1(s) = k_{11} * PE + k_{12} * CA + k_{13} * LE + k_{14} * ES + k_{15} * EQ \quad (1)$$

Technical preferences require different student's skills that relate more to the achieved results during the first part of the course, than to the ability to communicate or present. Therefore, we put the results of the exam to be on the top, followed by the quality of the written essay. Then comes the success of the laboratory exercises, and in the end we put the class attendance and the emotional intelligence as having the least influence on the technical feature. This order implies different values of the coefficients in Equation 2 as follows: $k_{21} = 0.30$, $k_{22} = 0.15$, $k_{23} = 0.20$, $k_{24} = 0.25$ and $k_{25} = 0.10$.

$$v_2(s) = k_{21} * PE + k_{22} * CA + k_{23} * LE + k_{24} * ES + k_{25} * EQ \quad (2)$$

Thus, the i^{th} student preferences profile is a vector $p_i = (v_1, v_2)$ containing the numerical values for each preference obtained from the valuation functions.

2.4.5 The allocation

Let's represent the fair division as a tuple $F = \{S, T, P\}$ where:

$S = \{s_1, ..., s_5\}$ is a set of students within a team;

$T = \{t_1, ..., t_5\}$ is a set of project's subtasks, and

$P = \{p_1, ..., p_5\}$ is a set of preferences of each of the five students in the team.

An allocation for F is mapping the students' preferences P into subtasks T, $f : P \longrightarrow T$, and for every $t \in T$ and $j \neq i$ such that $f(i) \cap f(j) = \emptyset$ and f is a complete allocation. That means that it is not possible to give the same subtask to different students at the same time and all the subtasks must be allocated.

We already ranked the subtasks by their importance for the project success. Considering the students preferences, we do the mapping by performing two types of ranking. The first ranking considers the responsibility preference value, v1. Once the manager is distinguished, the rest of the students are ranked by their technical preference values, v2, and are mapped into the matching subtask.

3 EXPERIMENTS AND RESULTS

To test our procedure we considered real data collected from 35 students enrolled in the Intelligent Systems course. Due to ethical reasons, instead of the students real names we use ordinal numbers.

Table 1 shows the results achieved in the first part of the course and the preferences derived from the results.

Table 1. Students achievements and their preferences.

Student	PE	CA	LE	ES	EQ	Responsibility	Technical
1	100	45	90	100	50	71.00	84.75
2	75	75	100	100	70	68.75	78.75
3	100	100	100	100	80	96.00	98.00
4	95	75	100	100	60	83.75	90.75
5	100	52	100	80	50	73.76	82.80
6	100	100	100	100	55	91.00	95.50
7	98	80	90	70	50	78.20	81.90
8	95	70	80	100	50	75.25	85.00
9	100	90	100	100	55	88.00	94.00
10	100	70	100	100	85	88.00	94.00
11	100	100	100	100	70	94.00	97.00
12	100	100	100	100	50	90.00	95.00
13	100	100	90	100	55	88.50	93.50
14	90	74	80	100	50	75.70	84.10
15	95	30	60	100	40	56.25	74.00
16	80	85	80	85	30	72.00	77.00
17	100	100	100	100	95	99.00	99.50
18	95	80	90	100	70	84.75	90.5
19	75	62	90	85	50	70.85	76.05
20	55	55	70	100	40	60.25	67.75
21	95	0	90	90	50	55.75	74.00
22	100	100	90	100	50	87.50	93.00
23	90	50	60	95	50	63.00	75.25
24	100	60	90	100	50	75.50	87.00
25	80	75	80	100	70	78.50	83.25
26	100	100	100	100	90	98.00	99.00

27	100	83	90	80	50	80.40	85.45
28	75	50	100	100	75	76.25	82.50
29	100	51	40	100	50	60.30	75.65
30	100	77	80	100	50	78.10	87.55
31	100	100	100	100	50	90.00	95.00
32	100	55	90	100	50	74.00	86.25
33	75	50	70	100	50	63.75	74.00
34	100	67	80	80	40	71.10	80.05
35	100	82	70	100	50	77.10	86.30

The responsibility and the technical scores are calculated by using the appropriate valuation functions as described in the *Modelling the preference* part in the 2.3.2 section. Having 35 students in total, we needed to make 7 teams with 5 students each, and assign one of the 5 tasks. The first step is to determine the managers of the teams. This is achieved by sorting the responsibility scores of the students. Since there are 7 teams, we take the 7 highest scores and assign the "management" task to those students. Removing the managers from the list, we have 28 students left for the rest of the tasks. All the 4 tasks that remain require technical skills. Therefore, we sort the students according to their technical scores, and assign to the top five "data analysis" task, then to the following five the "data processing task" and so on until each student is assigned a task, i.e. receives a portion. Having applied the approach, we obtained a distribution as shown in Table 2.

Table 2. Students tasks assignments.

Team	MG	DA	DP	RW	DPP
1	17	9	24	25	19
2	26	10	35	5	29
3	3	13	32	28	23
4	11	22	27	7	33
5	6	4	8	34	15
6	12	18	1	2	21
7	31	30	14	16	20

Having achieved the distribution of students, we need to check whether our allocation corresponds to their personal subtasks preference.

To check and compare each student's personal preference for a subtask we provided a survey among the students. We compared each student's personal preference for a subtask with the actual allocated subtask. 31.81% of the student's personal preferences matched with the allocated subtasks and 68.18% showed different preference compared to the assigned subtask. Providing more detailed analysis, we found that the students preferences that matched the assigned subtasks refer to: management, data analysis and data processing; whereas the students that were assigned either related work, or data preprocessing were dissatisfied. However, they have chosen another subtask that also belongs to the technical skills subtasks. This leads to a conclusion that the students are not willing to work on a subtask that is less valued. The overall analysis of the student's real preferences showed the preference choices as presented in Table 3:

Table 3. Students real preferences.

Subtask	Preferred by
Management	41.18%
Data analysis	32.35%
Data processing	26.47%
Related work	0%
Data preprocessing	0%

Considering these results, we can conclude that the students preferences are biased to the value for each subtask and not to their personal evaluation of his/her skills.

4 CONCLUSION

This research is about discrete subtasks fair division within student projects - a problem that cannot be solved by a standard cake-cutting approach. We deal with students that have the affinity either to minimize the required effort or to delegate the subtask to the other members of the team. Thus, we wanted to determine the students' affinities and provide a fair subtasks division in each team in order to achieve a best student performance in a given project.

We created an original methodology that analyzes student's activities during the first half of the course and suggests the most suitable subtask assignment. In order to have a fully automated procedure, we introduced a human factor for rating the discrete subtasks only at the beginning of the process. The subtasks are assigned to the agents according to one-to-one relationship. The assignment is obtained by a valuation function that maps the student's performance into the most suitable subtask.

This methodology has enabled finishing the projects in a more efficient way compared to the traditional approach - producing structured documentation and source codes on time and clear distinction of each student participation in the assigned project.

We performed a survey of the student's real preferences and analyzed the obtained results. The student's choice clearly supported the distinction between the Management subtask and the technical ones and it corresponds with our methodology of the subtask's division, but all the students has chosen only the highly valued technical subtasks. This outcome leads us to a conclusion that we should reduce the technical subtasks only to Data analysis and Data processing, assuming that Related work and Data preprocessing will be included as their integral part.

ACKNOWLEDGEMENTS

This work is financially supported by National project "Contribution to inclusive education of students with Down Sindrom", FCSE, 2019-2020.

REFERENCES

[1] S. Bouveret, J. Lang, "Efficiency and envy-freeness in fair division of indivisible goods: Logical representation and complexity", *Journal of Artificial Intelligence Research*, pp. 525–564, 2008.

[2] H. Moulin, *Fair division and collective welfare*. MIT press, 2004.

[3] S. J. Brams, A. D. Taylor, *Fair Division: From cake-cutting to dispute resolution*. Cambridge University Press, 1996.

[4] T. Burns, E. Roszkowska, N. M. des Johansson, "Distributive justice: From steinhaus, knaster, and banach to elster and rawlsthe perspective of sociological game theory", *Studies in Logic, Grammar and Rhetoric,* 37 (1), pp.11–38, 2014.

[5] Y. J. Cohler, J. K. Lai, D. C. Parkes, A. D. Procaccia, "Optimal envy-free cake cutting", in: *AAAI*, 2011.

[6] S. J. Brams, M. A. Jones, C. Klamler, "N-person cake-cutting: There may be no perfect division", *The American Mathematical Monthly* 120 (1), pp. 35–47, 2013.

[7] R. Simpson, M. Logan, P. Dolan, *Fair division algorithms*, 2013.

[8] R. Vetschera, D. M. Kilgour, "Strategic behavior in contested-pile methods for fair division of indivisible items", *Group Decision and Negotiation*, 22 (2) pp. 299–319, 2013.

[9] J. Rothe, I. Rothe, *Economics and Computation: An Introduction to Algorithmic Game Theory, Computational Social Choice, and Fair Divisio*. Springer, 2015.

[10] H. Aziz, S. Gaspers, S. Mackenzie, T. Walsh, "Fair assignment of indivisible objects under ordinal preferences", in: *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems, International Foundation for Autonomous Agents and Multiagent Systems*, pp. 1305–1312, 2014.

[11] S. J. Brams, D. M. Kilgour, C. Klamler, "The undercut procedure: an algorithm for the envy-free division of indivisible items", *Social Choice and Welfare*, 39 (2-3), pp. 615–631, 2012.

[12] S. J. Brams, M. Kilgour, C. Klamler, “Two-person fair division of indivisible items: An efficient, envy-free algorithm”, *Notices of the AMS,* 61 (2) pp.130–141, 2014.

[13] S. Bouveret, M. Lemaˆıtre, “Characterizing conflicts in fair division of indivisible goods using a scale of criteria”, in: *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems, International Foundation for Autonomous Agents and Multiagent Systems*, pp. 1321–1328, 2014.

[14] N. Dupuis-Roy, F. Gosselin, “An empirical evaluation of fair-division algo- rithms”, in: *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, Citeseer, pp. 2681–2686, 2009.

[15] J. D. Porras-Alvarado, Z. Han, Z. Zhang, et al., “A fair division approach to performance-based cross-asset resource allocation”, in: *9th International Conference on Managing Pavement Assets*, 2015.

[16] R. Zemel, Y. Wu, K. Swersky, T. Pitassi, C. Dwork, “Learning fair representations”, in: *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, pp. 325–333, 2013.

[17] B. T. Luong, S. Ruggieri, F. Turini, “k-nn as an implementation of situation testing for discrimination discovery and prevention”, in: *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM*, pp. 502–510, 2011.

[18] T. Kamishima, S. Akaho, J. Sakuma, “Fairness-aware learning through regularization approach”, in: *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on, IEEE*, pp. 643–650, 2011.

[19] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, R. Zemel, “Fairness through awareness”, in: *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference, ACM*, pp. 214–226, 2012.

[20] U. Endriss, N. Maudet, F. Sadri, F. Toni, “*Negotiating socially optimal allocations of resources*”, J. Artif. Intell. Res.(JAIR) 25, pp.315–348, 2006.

[21] M. Simjanoska, M. Gusev, S. Ristov, A. M. Bogdanova, “Intelligent student profiling for predicting e-assessment outcomes”, in: *Global Engineering Education Conference (EDUCON), 2014 IEEE*, pp.616–622, 2014.

[22] M. Simjanoska, S. Ristov, M. Gusev, “Generative modelling and classification of students’ e-learning and e-assessment results”, in: *Computational Intelligence, Communication Systems and Networks (CICSyN), 2014 Sixth International Conference on, IEEE*, pp.12–17, 2014.

[23] M. Simjanoska, M. Gusev, A. M. Bogdanova, “Intelligent modelling for predicting students’ final grades”, in: *Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014 37th International Convention on, IEEE*, pp. 1216–1221, 2014.

[24] G. Weiss, *Multiagent systems (intelligent robotics and autonomous agents series.* The MIT Press, 2 edition, 2013.