Science Journal of Business and Management
2015; 3(1-1): 73-79
Published online January 27, 2015 (http://www.sciencepublishinggroup.com/j/sjbm)
doi: 10.11648/j.sjbm.s.2015030101.22
ISSN: 2331-0626 (Print); ISSN: 2331-0634 (Online)

How matching algorithms can bring forth more effective decisions in situations with information deficiency

Péter Szikora

Óbuda University, Keleti Faculty of Business and Management, Budapest, Hungary

Email address:
szikora.peter@kgk.uni-obuda.hu

To cite this article:
Péter Szikora. How Matching Algorithms Can Bring Forth More Effective Decisions in Situations with Information Deficiency. *Science Journal of Business and Management.* Special Issue: The Role of Knowledge and Management's Tasks in the Companies.
Vol. 3, No. 1-1, 2015, pp. 73-79. doi: 10.11648/j.sjbm.s.2015030101.22

Abstract: University life is a chain of decisions. One of the most important parts of the decision as a process is the gathering and analysis of information, since the more information is available in case of a decision; the better one can define the options for the action, as well as their assessment. In most of the cases we simply don't have all/enough information, hence we make suboptimal decisions. Even in these cases, matching theory can offer a stable, optimal solution. Matching algorithms are one of the most important mathematical as well as economical approaches of the 21. century. Numerous university problems might be solved with the help of them. Nevertheless, although we very often apply some kinds of matching algorithms for handling decision situations, we are seldom aware of these algorithms which are most of the time ineffective. Present paper aims at proving that the conscious use of matching algorithms is not only for mathematicians, since their inner logic is easy to capture, and with the help of them the efficiency of the decision and the satisfaction of those involved in the situation may largely be improved.

Keywords: Matching Theory, Game Theory, Information, Knowing, School

1. Introduction

Life is a chain of decisions. In a broader sense decisions are nothing more than solving problems. Problems on the other hand might be defined as situations where our current status and the desired/targeted state are different, and we do not know the way leading from one to the other. We can talk about problem solving, when our actual status can be turned into the desired target [26] Decisions, when made to solve problems are rarely solitary choices; usually they should be interpreted as processes, with the first step being the recognition of the problem and the last the action resulting from the selection of one possible implementation. Continuous feedback is a necessary element of this process.

Students will also come across several problems in their university life, where decisions can be made with the help of matching algorithms. Such problems could be the application and admission into higher educational institutes, application for subjects and courses at or any decision where a distinct sets need to be paired with each other. In all such situations information is necessary to make a decision. In fact in many cases not only information but also knowledge is required.

2. Information, Knowledge or the Lack of Them

One of the most important parts of the decision as a process is the gathering and analysis of information, since the more information is available in case of a decision; the better one can define the options for the action, as well as their assessment. Complete information can never be achieved, so all decisions' inherent feature is the lack of information, and thus uncertainty is a given condition.

There is a big difference between knowledge and information, since information is simply a data or a sequence of data with a specific meaning, knowledge enables its owner to act. [26]

According to [20] there are two kinds of knowledge, tacit and explicit. Hence it resembles an iceberg where tacit elements, the experiences and the understanding of the situation at hand are the part that is under water, and explicit elements which can easily be articulated and transmitted are above the water level. The tacit and explicit forms of knowledge are not necessarily isolated from each other.

Hidden tacit knowledge through an upward spiral can be altered and become descriptive. The this way coded knowledge might on the other hand influence the behavior of the organizational members and be embedded into their habits, processes and knowledge.

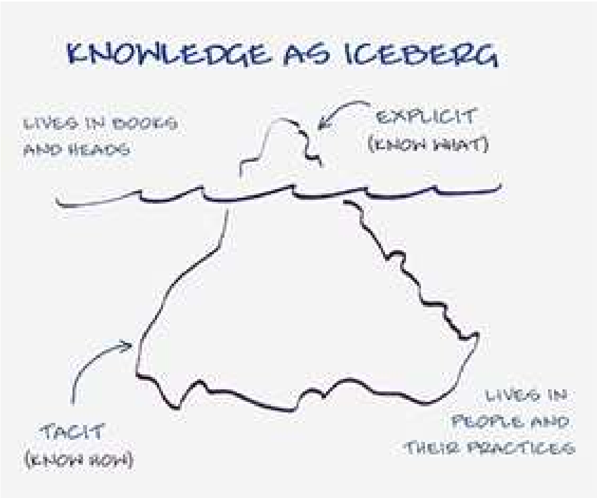

Figure 1. *Knowledge as iceberg by [19]*

Data become information, when they are provided a context and a form of conclusion is necessary in order to create intelligence. Intelligence is a capability that enables use in unexpected situations to use previously not experienced/known solutions. If we believe in these solutions and regard them as certainties knowledge is created, and by being able to connect and relate them intelligence is created. The hierarchy model of information with the links of various levels is displayed in figure 2.

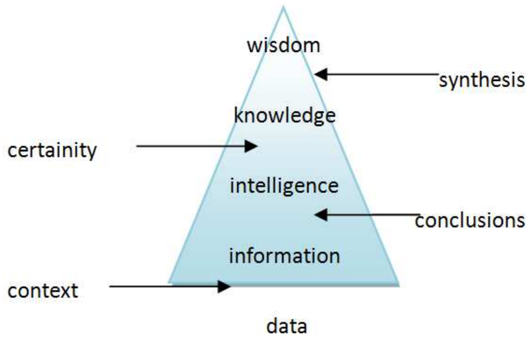

Figure 2. *Hierarchy model of information by [12]*

Accordingly in situations where students are supposed to decide, they would require knowledge, however most of the time they are even lacking basic information. What information is required for them to be able to decide for their own good? The most basic information should be the logic of the matching algorithm that is being used in the university systems (the most prevalent matching algorithms will be presented in this paper later on).The second information that is most of the time also missing is about their competitors' (other students) choices.

As an example let's see the process of course election at the university. Most of the time students pick their courses with the help of a simple greedy algorithm. The basic logic of this algorithm is that students are allowed to pick their courses in the order of their arrival, and those who once get into a course cannot be kicked out of it. However, matching theory can provide more effective means for matching as well, where sheer luck is not the sole factor in students landing in their desired courses. If they would be aware of the algorithm used, they could apply this knowledge in their selection of courses and point out their preferences or – if they want to use some tactics - the modified version of it. Nevertheless the second step is inevitable. Namely to gather information and based on it knowledge about their competitors' choices.

2.1. What is the Ground Cause of Information Deficiency

Our life is much more complicated than what we could perfectly simulate. The perfect model can never be prepared. We can never gather each and every piece of the necessary information, since there is not enough time for it, or the information is not accessible for us, or we do not have the financial resources to be able to obtain them. What is more, even if we could do so, the amount of such information would be so excessive that we would not be able to cope with them or process them. With the help of the internet an unlimited amount of information is accessible for us. However, even if we would use every minute of our 24 hours daily could we not even process the one millionth part of all the resources.

There are other limits to information processing. A human is able to perceive and process at most 7+/-2 digits of information at the same time. This number is the so-called magical number of [18]. We are able to forward these pieces of information from our short term memory to the long term repository; however, in order to do so, we have to connect them to another, already stored one or have to create a novel scheme. It is already proved that those who possess better schemes of the things in question are able to formulate better conclusions [26]. Accordingly, one of the biggest deficiencies of information processing is the lack of applicable schemes or the lack of ability to create such.

A further limit to information processing can be the cognitive limits of humans on attaining and using information. According to [17] the following cognitive limits occur when attaining new information:

- Deficiency of attention – limits already addressed in the previous paragraphs, namely humans are not able to obtain every necessary information, since time and resources are limited.
- Deficiency of memory – although we constantly perceive a wide variety of information, we are not able to remember most of them.
- Deficiency of comprehension – Obtained information has to be comprehended which is often impeded. New information is connected to the already existing ones poorly and if they contradict the already existing ones they are often disregarded.
- Deficiency of communication – not every human is

capable of transferring information to others. Each human uses his/her own unique schemes, so information communicated by one might be wrongly decoded, misunderstood by another. Hence information has to be communicated in different ways when addressing different people.

- Zoltayné added some other aspects to this list saying that [26]:
- We are searching for a structure, a scheme when obtaining new information where there might not be any or one that is totally unrelated to the one what we have found.
- Various people use various search mechanisms, what is more, the search mechanism for the same individual might change from situation to situation.
- People tolerate uncertainty poorly; hence they treat probabilities bad as well.
- It is also a prevalent situation to possess various bits and pieces of information but being devoid of the purpose and our aim in connection with them.

Hence the theory of problem-solving and decision-making has created the notion of heuristics, which is, a thumb rule to simplify strategies, to assist the creation of certain values and the resolution of problems (March, 2000) Upon decision, in order to clarify our own position we make simplifications and use assumptions and approximations in order to deal with the excessive information at hand. Hence heuristics simplify complex situations and provide aid in solving complex problems fast and effectively, however, we have to be aware of the fact that they might also distort our rational decisions. Since heuristics are traditionally applied in cases of uncertainty, when not applied carefully, problems are not solved, but multiplied and intensified by them.

Tversky and Kahneman enumerated the following basic heuristics [13]:

- Representativity: while disregarding the size of the sample the decision maker makes his/her judgment and assumes the probability of A and B choices on the basis of the choices' similarity.
- Accessibility: the probability of a given event is evaluated on the basis how easy it is for the individual to recall similar events – when the number of recalled similar events is bigger the assumed probability of the choice is also bigger.
- Anchoring and adjustment – the decision maker makes his/her judgements starting from a given point, assumption, and adjusts the perception of additional information according to the basic assumption.

These basic heuristics has been further developed by Max Bazerman [5]:

- Easy recollectibility
- Availability
- Apparent correlation
- Ignorance of prior probabilities
- Insensitivity to sample size
- Accidental misinterpretation
- Returning to the average
- Erroneous conclusions from the coincidence
- Insufficient adjustment
- Conjunctive and disjunctive events
- Excessive self-confidence
- The confirmation trap
- Curse of hindsight and ex post knowledge

The application of these (or any heuristics) - although they make decisions more easy - holds dangers that have to be kept in mind for the sake of satisfactory decisions. [23]

3. Matching Theory

In order to make a good decision, individuals do not only need information about the situation, and the other party's preferences, but also have to possess basic information about the logic of the matching algorithm that is being used.

Matching theory became one of the well-known research areas of mathematics as well as economics research in the 20th century. For a long time we have been searching for the means how to efficiently match sets with each other. We just have to consider mathematical problems such as linear programming or transportation and assignment of tasks. The disadvantage of these is the aim being the optimisation, the maximisation of the total profit.In case of matching theories on the other hand the purpose is to maximise the participants' individual profit.

The first article on matching theory has been created by Shapley and Gale [9] , introducing the problem of matching through the emergence of marriage relations, and proving the (existence of) stable solutions. With the help of a simple algorithm they have proved the existence of a stable solution, in case the preferences of both males and females are given. It is important to note that they did not only create pairs, but created so called stable pairs (marriages), where there are no blocking pairs offering benefits for both of the partners for leaving their present partners at the very same time. Owing to the success of this algorithm matching theories are being used with residency problems - placement of resident physicians to various institutions - National Resident Matching Program, NRMP - ([21]; [9]), in the secondary and tertiary educational admission [1][3][11][21][22][6][7][8][14][15]. In the '90s the starting point of the NRMP has been changed. Now the algorithm is not being applied from the institutions' point of view resulting in the worst possible outcome for the students but from the students' point of view.

If we suppose that institutes (hospitals and educational institutes) are not behaving like strategic partners, namely they do not use tactics; the outcome of this algorithm may even result in the revelation of the agents' true preferences.

The purpose of present study is to prove that even in the everyday university life there exist various problems that can be solved (more) effectively with the help of matching algorithms.

3.1. Matching Theory Models

The purpose of the matching is to pair the elements of two disjunctive sets according to their preference orders. The first such matching problem for which a natural algorithm has been initiated by [9] was the marriage problem. If the set of peaks can be distributed into two subsets (e.g. males and females), where edges are exclusively between the two subsets the graph can be considered bipartite graph. There are matching problems, where the creation of two subsets is not inevitable, such as the room-mate problem, where each player can be paired with every other player.

In case of the marriage problem there exist two disjunctive sets, let's label them with F and M; F standing for females and M for males. The elements of the sets can be represented with f and m characters. Than we can state that

$$F \cap M = \emptyset$$

and

$$F = \{f_1, f_2, \dots f_n\},$$

and

$$M = \{m_1, m_2, \dots m_p\},$$

if there are p number of males and n number of females. The preference order for males and females also has to be determined. The preference order of females is:

$$P(f_1) = m_3, m_1, m_2 \dots,$$

while that of males:

$$P(m_1) = f_3, f_1, f_2 \dots,$$

This way the matching possibilities can be described with the (F, M, P) relation.

Marriage takes place between a male and a female if they are on each other's' reference list. The main point of the algorithm to provide the best possible match regarding the person's preferences the algorithm runs for. If the other party does not decline the proposal (because he/she has not been proposed to yet, or the previous partner is worse than the present according to the individual's preferences) the two will be considered a momentary match. In each round, those who have been declined will search for the next best partner according to their preferences. The last step is for the temporary partners to become final matches. The matching created this way will be stable.

This is a matching since each male has at most one female partner, and in return each female has also top one partner. In order to prove the stability of the system let's regard a male and a female who are not each other's partners in the final round. This could have happened because of two reasons. Either the male has proposed to the female, but has been rejected, or the male has never proposed. If the male has been cast off, at one point of the algorithm that means the female at that point of the time had a better proposal. However, since each proposal accepted has been approved if and only if it had been better than the previous, it is sure, that the female ended up with a better proposal than the male in our example. If the male has never even proposed to the female that means, that throughout the matching he has been courting females, who were higher on his preference list, and at the final stage of the algorithm ended up with a partner, who he preferred to the female in our example.

3.2. Matching Theory Algorithms

Greedy algorithm

In the random or sequential application of dictatorship, the agents of one set are ordered (randomly, or by drawing lots), and the current agent as if he/she were a dictator may choose between the existing (remaining) options. Thus, the algorithm does not take into consideration the other party's, (the agents of the other set) preferences.

To better understand the algorithm let's consider each agent of the first (decisive) set one by one:

- if he/she can find an option, that is preferred, the agent is matched with the most preferred option – one slot becomes occupied
- if we cannot find an empty slot (option), that is preferred, the agent remains unmatched.

Gale-Shapley algorithm

In addition to the Greedy algorithm not being a stable matching algorithm, in case of some problems, like school admission processes another problem also arises. Namely that each agent in both sets has its own preference order that is, or might be different from the others' order. Hence a school admission algorithm has to take into account and consideration the preferences of the educational institutes beside those of the students. [4] and [1] pointed out that the Gale-Shapley algorithm is not only able to handle the case of double preferences but can deal with extra rules, such as controlled choice, where certain limitations are in place to diminish the ethnic and racial segregation. In order to have a small insight into the logic of the Gale Shapley algorithm lets label the first set of agents students, and the other one schools.

1 In the first step each student applies for the school with the highest preference ranking.
2 When more students designate the same school than as many places it has to offer, the schools select the students with the highest preference points, and reject all other students.
3 All students rejected in the first round apply for the school with the next highest preference ranking and the second step is repeated again. If a student with higher preference ranking occurs than the previously accepted ones, he/she can take the place of the ones accepted in the first round if there is not enough place for each of them.
4 The ones without an admission apply for the school with the next highest preference ranking and the third step is repeated again.
5 Step 4 is repeated till each student finds a suitable school or gets to the end of his/her own preference list.

Boston mechanism

This mechanism has been used in Boston and some other towns between 1999 and 2005 for simulating and solving the problem of high school admissions [11][2][1]. For a deeper understanding of the algorithm let's suppose the two sets being that of students and that of schools the students apply for – the same as in the original application.

1 In the first step each student applies for the school with the highest preference ranking.
2 When more students designate the same school than as many places it has to offer, the schools select the students with the highest preference points, and reject all other students. Students already selected cannot use their right to attend the school because a new student applies for the same school with higher preference points.
3 All students rejected in the first round apply for the school with the next highest preference ranking and the second step is repeated again.
4 Step 3 is repeated till each student finds a suitable school or gets to the end of his/her own preference list.

The mechanism is Pareto-optimal on the basis of the given preferences. The biggest deficiency of the system is that agents have to use tactics, hence it is of high risk to apply for an oversubscribed school, since if one does not succeed in getting in he/she might end up with his second and third choice schools being already full before the second round starts [10]

The Colombus method

The method used in Colombus resembles the system that has been used before the central allotment was introduced into the resident allocation problem. The main difference is that those already accepted to a place are not being considered as a part of the set any more and thus will not be eligible for a better proposal that could destabilise the system and cause the problem of legitimate envy. The algorithm that has been applied in Clombus City [1] is as follows:

1 Each student may designate (apply for) 3 different schools.
2 Certain schools guarantee places for those who live in their districts, for the remaining places the applicants are ordered by random choice.
3 The free places are offered to students who are the top ranking ones on the school's preference list.
4 The students have to make a decision concerning the proposal within three days. If they accept it, they will be matched with the school at hand and are removed from the set of applicants and will not be offered any other options.
5 The offers declined will be made to students on waiting list who are subsequent to the previously proposed ones and step number 4 is repeated.

Top trading cycles method

The main point of the algorithm proposed by [1] is that students selected by schools can trade their places among themselves. Its advantage is that it is in the students' true interest to disclose their real preferences, so the algorithm provides remedy for most of the concerns raised by previously introduced methods. The algorithm runs as follows:

1 Each student and school identifies what / who is ranked in the first place. As the number of participants is limited, there is an s_1., C_i, s_2., ..., $C_k$cycle where s_i. prefers C_i, who on the other hand prefers$S_{i+1}$and at the end of the cycle C_k prefers s_1. Each student and each school belongs to not more than one cycle. Every student who belongs to such a cycle will be admitted by the school designated by the student. With this the student is removed from the system, and the school ends up with one less free place remaining. If all places of a certain school are matched, the school will be removed from the system, so the remaining students can no longer designate it as their favorite.
2 In every further step all remaining students and schools are involved, apart from this the step follows the same course as the first one. Participant name their preferences and students belonging to a certain circle are admitted to the school that they name (and belongs to their circle).

The algorithm ends when there are no more students who have not been admitted to any schools. Accordingly, since at least one student is admitted in each cycle, the number of steps cannot be more than the number of students.

A comparison of different algorithms can be seen in the following table.

Table 1. *Comparison between different matching's algorithm*

	Rewards Strategising?	Holds 'Place in Line'?	Strongest Driver:	Rewards True Choices?
Greedy	no	no	Firstchoice	notalways
Gale-Shapley	no	yes	priorities; any choice	yes
Boston	yes	no	Firstchoice	Notalways
Columbus	yes	no	Firstchoice	Not always
Top Trading Cycles	no	yes	top choies; trading	yes

As displayed in the table, there are algorithms that support the applicants to reveal their real preferences, while some algorithms foster using tactics. However, it is no coincidence that the Gale-Shapey algorithm is applied for various admission processes, be it secondary or tertiary education education, or the resident allocation problem.

3.3. Matchings in University Settings

As already presented before, there are many different algorithms and in various situations, different algorithms are used. Unfortunately, at our university besides the admission process the Greedy algorithm is used everywhere. For which not being a stable matching is not the only concern, as already expressed in the previous section. Another of its deficiencies is that with the help of it students cannot be ranked. Accordingly admission is rather a question of luck.

Unfortunately even at an international level, we know of very few situations where stable matching algorithms are used, these can be seen in the following table.

Table 2. *Matching schemes in Europe*

Country	Student admissions	Job market	Kidney exchange	Other applications
France		professor allocation		
Germany	higher education			
Hungary	secondary schools,higher education			
Israel				dormitories
Netherlands			The Dutch program	
Spain	higher education		The Spanish Program	
Turkey	higher education			
UK		SPA-SFAS,TIS		

Source: [16]

As it can be seen in table 2 there are very few places where matching algorithms are used, and if, greedy algorithms are run as a mean of matching. It would largely improve the efficiency of the processes and the satisfaction of those involved if a different (stable, more efficient) matching algorithm would be used.

4. Conclusions

In university life there are various situations where suboptimal decisions are made, hence the stakeholders' satisfaction is low. Students usually lack information necessary for an optimal decision. Sometimes the decision of joining a certain higher educational institute is already a bad decision, since the satisfaction, and hence their attendance of classes is low [24]

The causeof this deficiency is twofold. On the one hand it is hard (impossible) to attain information, on the other hand, even if one would succeed in acquiring all the necessary information the processing of them would require too much time, energy and resources. Matching algorithms could provide an easy to use and efficient solution to the problem at hand. The total satisfaction could be increased by 22 percent with the help of Gale-Shapley, 21 percent with the help of Boston algorithm [25] As described in present article, the Columbus and TTC methods might produce even better results.

Unfortunately, in university life matching is determined mostly by luck, and most of the time the first come first served greedy algorithm is being applied. The preferences of neither of the involved parties are taken into consideration. It would largely improve the efficiency of the decisions if a stable, more efficient matching algorithm would be used where preferences as well as individual and total profit is taken into consideration.

Acknowledgements

First of all, I would like to thank my colleague Kornélia Lazányi for all of her help in this work. Also I would like express my thanks to my Advisor László Á. Kóczy who played a decisive role.

References

[1] Abdulkadiroğlu, A.–Sönmez, T. , "School choice: A mechanism design approach". American Economic Review, 93. pp. 729–747, 2003

[2] Abdulkadiroğlu, A.–Pathak, P. A.–Roth, A. E.–Sönmez, T. "The Boston public school match". American Economic Review, 95. pp.368–371, 2005

[3] Abdulkadiroğlu, A.,Che Y, Yasuda Y., Expanding „Choice in School". Choice Economic Research initiatives at Duke, 2008

[4] Balinski , M.–Sönmez, T.: "A tale of two mechanisms: Student placement". Journal of Economic Theory, 84. pp 73–94., 1999

[5] Bazerman, M. H.: "Judgment in managerial decision making", John Wiley & sons, New York, 1990

[6] Bíró Péter, "Stabil párosítási modellek és ezeken alapuló központi párosító programok". Szigma, 37. pp. 153–175, 2006

[7] Biró Péter, 2008, "Student Admissions in Hungary as Gale and Shapley Envisaged". Technical Report TR-2008-291, University of Glasgow, Department of Computing Science, Glasgow.

[8] Biró Péter–Fleiner Tamás–Irving, R.–Manlove, D., "The College Admissions problem with lower and common quotas." DCS Technical Report TR-2009-303, University of Glasgow, Department of Computing Science, Glasgow. , 2009

[9] Gale D., Shapley L. S., "College admissions and stability of marriage". American Mathematical Monthly 69: pp. 9-15,1962

[10] Glazerman, S.–Meyer, R. H.: "Public school choice in Minneapolis". Downes, T. A.–Testa, W. A. (in edit.), Midwest approaches to school reform. Federal Reserve Bank of Chicago, pp. 110–126. 1994

[11] Ergin, H.–Sönmez, T.,"Games of school choice under the boston mechanism". Journal of Public Economics, 90. pp. 215–237, 2006

[12] Haeckel, S. H.Presentation to the information planning Steering Group, Marketing science Institute, Cambridge, MA, 1987

[13] Kahneman, Daniel and Tversky, Amos: „Prospect Theory: An Analyis of Decision under Risk” Econometrica, Vol. 47, No. 2, 1979

[14] Kóczy Á. László, “Központi felvételi rendszerek. Taktikázás és stabilitás”, Közgazdasági Szemle, LVI. évf.,pp. 422–442. , 2009

[15] Kóczy Á. László, “A magyarországi felvételi rendszerek sajátosságai Magyarországon”, Közgazdasági Szemle, LVII. évf.,pp. 142–164. , 2010

[16] Kóczy Á. László, “Matching schemes in Europe”. http://econ.core.hu/english/res/game_app.html , Downloaded: 2014 november 25.

[17] March G., James: “Bevezetés a döntéshozatalba” Panem kiadó, Budapest, 2000

[18] Miller, George A.: “The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information”, The Psychological Review, 1956, vol. 63, pp. 81-97

[19] Paterson,R,”JSB on Education Explicit Versus Tacit Knowledge” http://smartpei.typepad.com/robert_patersons_weblog/2004/04/jsb_on_educatio.html, downloaded 2014. november .25.

[20] Polanyi, Michael,“The Tacit Dimension”. Garden City: Doubleday and Company, 1966

[21] Roth A. E., “The evolution of the labor market for medical interns and residents: a case study in game theory”. Journal of Political Economy 6: pp. 991-1016, 1984

[22] Roth A. E., Peranson E.,“The redesign of the matching market for American physicians: some engineering aspects of economic design” The American Economic Review 89:pp. 748-752,1999

[23] Sterbenz, Tamás „Korlátozott racionalitás a sportmenedzseri döntésekben”, doktori (PhD) értekezés ,Nyugat-Magyarországi Egyetem, Sopron, 2007

[24] Szikora, Péter “Tanítás, mint kooperatív dinamikus játék“, In: Cser L, Herdon M (editor) Informatika a felsőoktatásban 2011 konferencia. 1140 p. Konferencia helye, ideje: Debrecen, Magyarország, 2011.08.24-2011.08.26. Debrecen: Debreceni Egyetem Informatikai Kar, 2011. pp. 947-954.

[25] Szikora, Péter “Allocating time-bound tasks – an application of matching theory”SEFBIS Journal 2015. in press

[26] Zoltayné Paprika Zita “Döntéselmélet”, Aliena kiadó, Budapest, 2005