Consensus Moderation System

Andrei TOMA
Academy of Economic Studies, Bucharest, Romania
andrei.toma@ie.ase.ro

The present paper formulates a consensus moderation system based on the negotiation of the actors involved. There are a series of steps in the moderation process, the first of which is constructing a front of Pareto optimal solutions. Since this in itself will likely not lead to consensus in a real life scenario, Kaldor-Hicks compromises are then detected. Compromises are recommended at every iteration of the negotiation process which can lead to a lengthy negotiation time, which is addressed by using a recommendation engine based on the previous behavior of the actor.
Keywords: *Pareto Optimality, Kaldor-Hicks Compromises, Consensus Model, Consensus Moderation*

1 Introduction

A consensus moderation system would prove a valuable tool in the administration of negotiation between a small number of actors. In this paper I propose an approach to constructing consensus and propose the conceptual architecture of a system integrating this approach.

There are several possible directions which such a system could explore, approaches which are briefly examined below.

Simple collaborative approaches are mechanisms created to facilitate the interaction of citizens. The construction of a common position which satisfies all the participants is not done by the system itself, as such a result is emergent from the interaction of the citizens themselves.

It is very difficult to limit this category to a series of tools since practically any collaborative mechanism can be used to construct a common opinion. For example, any wiki software could be used as a political tool with the singular restriction of the goal followed by the majority of its users.

As such, simple approaches rely more on the existence of pre-existing social mechanisms than on a technical solution. While they constitute useful tools in the creation of consensus, they could, at best, be used as support tools to a systemic approach.

Statistical collaborative approaches involve a statistical approach to the interactions between users. The simplest way to apply such an approach would be to construct an aggregate opinion formulated by statistical means, such as averaging the scores users associate to certain proposals/comments. While this type of approaches does have some merits and has the benefit of realistically reflecting a classical democratic mechanism, it will generally not be enough to create complex integration and guarantee the construction of consensus.

Approaches involving more evolved mechanisms try to construct a common opinion by including artificial intelligence mechanisms.

In the context of artificial consensus moderation, I will try to supplement the shortcomings of a simple system (which can be constructed to guarantee reaching a necessary, but not sufficient condition for consensus) through the addition of a recommendation system which will try to emulate the decision strategies of the actors involved.

2 Constructing consensus

Through consensus moderation I understand an approach which tends to accelerate the convergence of the ideas of a population of users through some type of automatic mechanism. Although total convergence itself through an automatic mechanism is unlikely, the more efficient the system is, the higher the chances for consensus are.

The construction of a consensus moderation system is attempted through an abstract approach which allows us to analyze the process of government as a system in which the relations between the actors involved can be analyzed and improved.
Since laws relating the information technology are dependent in their application by wide industry acceptance, the system will allow these actors to reach an acceptable middle ground which, in turn, will eliminate some of the misgiving the actors involved have.
If applied to larger entities proposed system will, on one hand, allow the citizens of a state to be directly involved in the construction of laws and, on the other hand, allow the state to rely on a solid democratic legitimacy. While the final purpose in this case is the one above, an intermediary system would involve actors to which a law is to be applied. For example, for the construction of a law concerning digital content provision, at least the major players on the market should be involved, alongside other important organizations, even if the ideal situation, in which all citizens have a distinct voice, is impossible.
A consensus moderation system would help the parties involved, whichever they may be, in constructing a legal text which reflects the interests of all participants through a series of interdependent compromises. Such a system would produce a legal text which would serve either to legitimate the efforts of the legislator or to show that the will of the people to whom the law is to be applied is opposite to the legal situation.
In trying to achieve the construction of artificial consensus, several approaches relating to meta-government have been studied. A combination of these approaches is the best solution for such a system.
Through the idea of consensus moderation we understand the construction of a network of compromises between the actors involved in the negotiation which allows these actors to reach a solution considered satisfactory by all of them.
The concept on which the idea of moderation is built is that it is possible that in a negotiation with a sufficiently large number of actors an optimal solution might exist that the actors themselves are not aware of. The ideal system would come to the aid of the participants through the construction and recommendation of this ideal compromise solution.
A moderation mechanism, as it is understood in this paper, represents two different concepts. On one hand, through “moderation mechanism” I understand the general moderation methodology by which aiding the construction of consensus is approached. Under this definition, the moderation mechanism consists of a series of preliminary steps, such as decomposition of proposals, followed by a number of algorithms which are to be run in order to ensure the best chance of consensus being reached.
On the other hand, by “moderation mechanism” I understand the actual algorithm by which the proposal space is stabilized, in the sense that the actors are determined to group themselves around common opinions.
There are a series of possible problems which must be taken into account when constructing a prototype system. Some of these will be addressed in the following sections, while some are the possible subject of future research, as they go beyond the scope of this paper.
The first problem the system will encounter is managing the issues the users raise in connection to a particular proposal. In order to maintain a higher chance for consensus to be reached, issues are not allowed to overlap. In a first phase each issue is discussed separately, before deciding to go into the compromise phase. The decision of splitting an issue is automated, although is origin is always one of the actors involved.
Another important problem is separating the proposals, issues and solutions which are made “jocandi causa”, in the sense that they do not address real problems and they are raised in order to create perturbations in the system. This remains a serious problem in

the event that such a system might be applied to small communities of regular citizens. In the study done in this paper, however, the actors are assumed to be large business entities which are more interested in reaching an acceptable compromise than in creating perturbations in the functioning of the system.
A more important problem, however, is that the system must offer a solution for the cases where compromises cannot be reached. Since the very problem of consensus is its blocking character and the fact that trying to reach consensus might result in no text being adopted, this is one of the main issues with which the system must be concerned.
While situations where consensus cannot be reached will still be possible, the objective of the system is to minimize the number of such occurrences and thus lead to a more effective negotiation process.
If the system is to be used as a mechanism for confirming/infirming the democratic legitimacy of a state body, the only effective way of testing it is having it used by a large enough number of citizens.
Since the system has its primary intended used for the negotiation between the state and large industry players, it can also be tested on a relatively small set of actors negotiating laws and regulations pertaining to a certain field.
For this case, the system can be tested either in the actual negotiation between actors represented by a human "spokesperson" or via a sufficiently defined set of agents simulating their behavior.

3 Conceptual architecture of the system

The actual system would have to integrate the steps described above, alongside classical collaborative mechanisms allowing for the effective communication between the parties involved. For the approach based on a predicted future behavior of the actors, the actions of the users would have to be stored in the system and used as input for the recommendation engine.
There are a series of steps which must be followed in the discussion of a proposal, steps which will be reflected in the sections below.
The flow of the system is reflected in the Figure 1 below which reflects the steps through which a proposal must pass in order for consensus to be reached. As we can see, there are a series of steps that are followed in the flow of a proposal through the system. First the proposal is registered by a proponent with the system. The next step is for proponents to raise issues.
If issues overlap, they must be decomposed in different, smaller issues, which will have the same proponents as the "parent" issue involved in the negotiation.
For a defined issue, there is a particular lifecycle of its negotiation in the system, which is composed by a series of iterations through the negotiation procedure. For each iteration, there is a series of phases through which the proposal must pass. First proponents can then propose various solutions they envision. After the solutions are registered with the system, they are then endorsed by the other proponents; each endorsement has an associated score, which can either be visible or invisible to the user.
If there are several solutions for a particular issue, the system then tries to eliminate the subset of the solution which does not satisfy the Pareto optimality criterion. If, after eliminating the dominated solutions, there are still multiple solutions, possible compromises are proposed to the proponents on different sides of the issue.

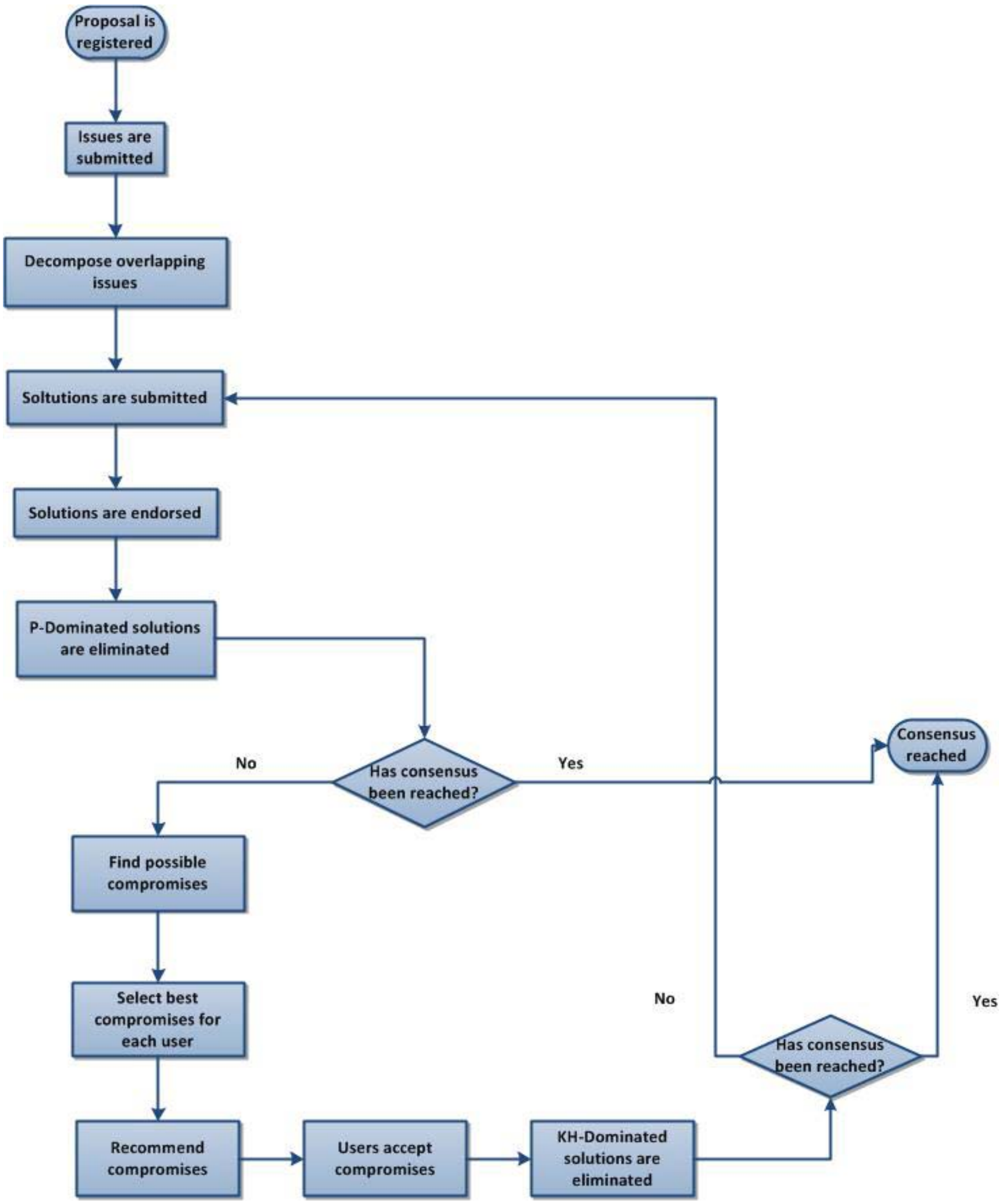

Fig. 1. Flow of the system

In order to accomplish this, the possible compromises which satisfy the Kaldor Hicks optimality criterion are selected. While all these could be presented as possible ways of reaching consensus to the proponents, it is better to select only the compromises which are more likely to be accepted by a proponent (a fact which is determined according to their prior behavior).

If the proponents accept the recommended compromises, the system removed the Kaldor Hicks dominated solutions.

At this point, the system either has a possible consensus, or the system must go through another negotiation iteration.

If consensus is possible, the proposal is said to be stable, providing it passes a final round of approval from the proponents.

4 Pareto optimality of solutions

The idea of using Pareto efficiency as a necessary, but not sufficient, criterion for the construction of consensus has been used before [1]. The principle that leads to the use of such an approach stems from the principle of Pareto optimality, since a dominated proposal would not have lead to greater opinion convergence.

Through opinion convergence I understand the degree to which the different parties

involved in the negotiations agree with certain proposal. The goal of optimizing convergence is to finally reach a single opinion which satisfies the parties and thus reach general consensus.
The concept of Pareto efficiency [9] [10] is an economic concept, but has also historically been used in social sciences. Pareto efficiency is a minimal notion of efficiency and does not necessarily result in a socially desirable distribution of resources: it makes no statement about equality, or the overall well-being of a society.
However, a Pareto inefficient system is to be avoided since if a reallocation of resources is likely to make some individuals' situation better without harming the others, it is clear that the situation of the whole population of agents can be improved.
Given an initial allocation of goods among a set of individuals, a change to a different allocation that makes at least one individual better off without making any other individual worse off is called a Pareto improvement.
An allocation is defined as "Pareto efficient" or "Pareto optimal" when no further Pareto improvements can be made (thus there are no obvious improvements to the situation of any participant).
If economic allocation in any system is not Pareto efficient, there is potential for a Pareto improvement (an increase in Pareto efficiency): through reallocation, improvements to at least one participant's well-being can be made without reducing any other participant's well-being.
Using the classical "butter vs. guns" example a Pareto efficient situation appears in the figure below where applying Pareto optimality leads to a production curve.
Point A is not efficient in production because you can produce more of either one or both goods (Butter and Guns) without producing less of the other. Thus, moving from A to D enables you to make one person better off without making anyone else worse off (rise in Pareto efficiency). Moving to point B from point A, however, is not Pareto efficient, as less butter is produced. Likewise, moving to point C from point A is not Pareto efficient, as fewer guns are produced. A point on the frontier curve with the same x or y coordinate will be Pareto efficient [6].

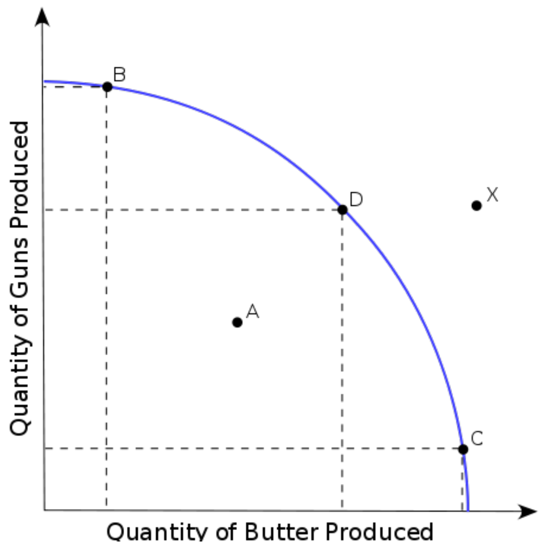

Fig. 2. Production-possibility frontier [2]

When discussing Pareto optimality, two concepts must be taken into account, that of a weak Pareto optimum and a strong Pareto optimum. A "weak Pareto optimum" (WPO) satisfies the condition of not being Pareto inferior to a different allocation of the resources, but the definition of a possible improvement differs from the case of a "strong Pareto optimum".
For a WPO to be reached, it is sufficient that no other allocation is strictly preferred by all other individuals. This might mean that there exists a different allocation that improves the situation of part of the individuals while leaving the situation of the others unchanged, but such an allocation is not covered in searching for a WPO. In other words, a WPO is reached when no reallocation exists that will cause all individuals to gain.
Weak Pareto-optimality is "weaker" than strong Pareto-optimality in the sense that the conditions for WPO status are "weaker" than those for SPO status: any allocation that can be considered an SPO will also qualify as a WPO, but a WPO allocation won't necessarily qualify as an SPO [6].
A "strong Pareto optimum", on the other hand, is reached when there are no other allocations which would allow some

individuals to gain while not deteriorating the situation of any other individuals. As such, a reallocation that leads to a SPO might be strictly preferred by some individuals (those whose situation is actually improved), but non strictly preferred by the rest (those whose situation is not worsened, but not improved).

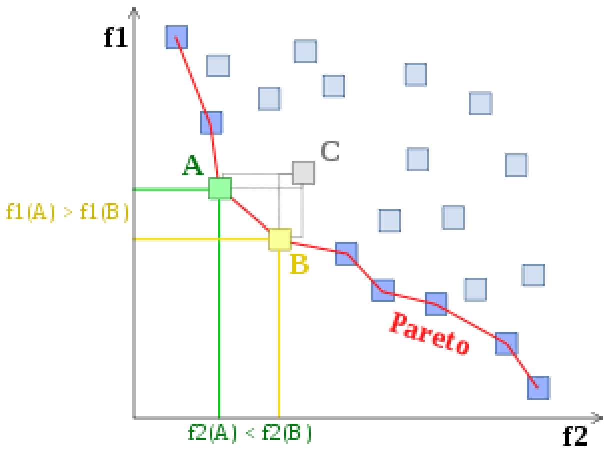

Fig. 3. Pareto frontier [3]

Figure 3 shows an example of a Pareto frontier. All the squares in the figure are possible choices with smaller values preferred to larger ones. As an example, moving the frontier to point C would not be preferable from a Pareto point of view as it is dominated by points A and B.

In order to obtain consensus in the negotiation of a particular issue, it is useful to bring the possible solutions to a state of Pareto optimality. This will mean that all participants obtain maximum possible satisfaction under the restriction that no dissenting participant's opinion can be ignored.

The system attempts to create Pareto optimality for each issue before trying to reconcile between issues that could not successfully be brought to consensus through this minimal criterion.

A possible solution S1 approved by a set of users U1 to an issue is dominated if there is another solution S2 that has been approved by a set of users U2 which is a superset of U1.

$$U1 = \{u \in U | pref(u, S1) = 1\}$$

$$U2 = \{u \in U | pref(u, S2) = 1\}$$

$$pref(u, S) = \begin{cases} 1, \text{if } u \text{ approves } S \\ 0, \text{otherwise} \end{cases}$$

$$U1 \subseteq U2$$

The steps in trying to achieve Pareto optimality in the solution space are presented in the algorithm below. These steps will be run for each generation of the algorithm.

Algorithm:
P1.Select solution **S1**
P2.Select solution **S2** form **S\{S1}**
P3.If **S1** dominates **S2** remove **S2** from the solutions

Fig. 4. Removing dominated solutions

Once dominated solutions have been eliminated, the system might reach a stable state, in which there is a dominant solution for each proposed issue. However, this will most likely not be the case, so the next step is to identify possible compromises. If these are accepted, opinion convergence will be increased in that there will be less contested issues within a proposal.

5 Constructing Kaldor Hicks compromises

Pareto efficient systems are hard to attain in social systems so usually Pareto efficiency is used as a baseline test for compromise based systems.

Since from the previous section the system is in the situation that it cannot lead the system to consensus, we must widen the search for possible solutions. In order to achieve this, one possible approach is using the Kaldor Hicks criterion which relaxes the conditions of searching for Pareto improvements in order to better apply to real life situation.

The Kaldor Hicks criterion has been widely applied in the field of welfare economics [4] to create actual policies when changes were required which could not be conceived as

Pareto reallocations. An important note however is that using the Kaldor Hicks criterion for policy construction is subject to controversy since there is no way of guaranteeing that the compensation the "losers" should gain is actually distributed to them.

I believe however, that in the situation that is covered in the consensus moderation system, the criterion can be applied while circumventing these limitations. For one, compensation is automated since all "transactions" happen inside the system. Another important element is that all actors are supposed to be aware of the way the system functions which lessens the probability of an actor accepting a compromise where it is not compensated.

Nevertheless, Kaldor Hicks improvements should be handled with care, as not all of them also lead to Pareto optimality. This can be seen in the figure below [5].

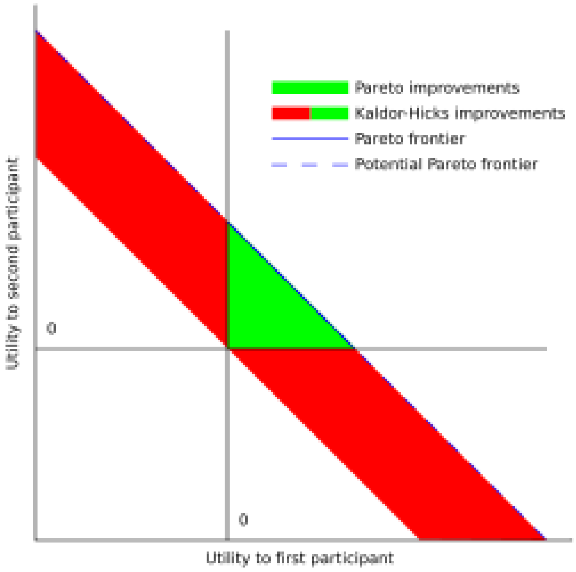

Fig. 5. Kaldor-Hicks improvements [5]

As can be seen from the figure above, represented by the green area, only a small subset of Kaldor Hicks improvements actually lead to a more Paretto efficient situation.

The red area represents possible compromises which lead to at least one party being worse off and thus do not meet the criteria of balance which I am trying to achieve in the system.

In order for an improvement to meet the Kaldor Hicks criterion, two conditions must be met:

That the party "winning" from the compromise can adequately compensate the parties "losing" in order for them to not be worse off after the redistribution

That the parties "losing" from the compromise could not have adequately compensated the winner

In effect, any redistribution which fits the Kaldor Hicks criterion must be unidirectional and not worsen the situation of any party. As stated above, redistributions must also be of such a nature as to ensure that compensation actually happens. This is because the criterion only requires that compensation be possible, not that it is actually played. This is less of a problem in an automated system where compensations happed without a party having the chance of actually not fulfilling their obligations.

Searching for Kaldor Hicks compromises is presented in the equations below, which attempt to formalize the construction of compromise between actors discussing multiple issues.

The equation below defines the issue set, which is represented as a set of solutions which the users have proposed for a particular issue

$$I_i = \{S_{i,j} | S_{i,j} \text{ is a proposed solution to } I_i\}$$

Apart from the issue set, we must also represent the set of the users involved in the negotiations. All variables noted *u* should be considered as belonging to the users set.

$$U = \{u_k | u_k \text{ is a user in the system}\}$$

In order to reflect the attitude of a user towards a solution proposed to a particular issue, we must introduce a preference function (*ipref*), which is defined below. There are two possible forms for the function, the binary form and the extended one used here. While a binary preference function is sufficient in order to reflect Pareto dominance of solutions, it would be impossible to reflect the "price" a compromise has for the "winner" of the compromise.

$$ipref(u_i, S_{i,j}) = \begin{cases} 0, if\ u_i\ does\ not\ endorse\ S_{i,j} \\ n > 0, the\ preference\ of\ u_i\ for\ S_{i,j} \end{cases}$$

We define the set of endorsing users for a particular solution as *E*. This content of each set is dependent both on the solution and the issue it is offered for. In defining *E*, it is only relevant that a particular set of users has manifested some degree of approval, but not the actual level of approval they provided.

$$E_{i,j} = \{u_k | ipref(u_k, S_{i,j}) > 0, u_k \in U\}$$

There is a connection between the sets of endorsing users and solution domination in the Pareto sense. The concept of solution domination only has relevance for solutions proposed for the same issue. A solution $S_{i,j}$ dominates another solution $S_{i,k}$ if the corresponding endorsing set for $S_{i,j}$ includes the corresponding endorsing set for $S_{i,k}$.

$$E_{i,j} \subset E_{i,k} \Rightarrow S_{i,j} \prec S_{i,k}$$

Apart from the set of users endorsing a solution, in order to construct a compromise, we must have the set of users who do not agree with the solution, even to a minimal extent. Hence their preference for the solution must be 0.

$$\overline{E}_{i,j} = \{u_k | ipref(u_k, S_{i,j}) = 0, u_k \in U\}$$

The next step is to actually detect possible Kaldor Hicks compromises, which will then be recommended by taking into account prior user behavior.

The set of Pareto dominant solutions for an issue is represented by solutions which are not dominated by any other solutions proposed for the same issue, as seen in the equation below.

$$P_i = \{S_{i,j} \in I_i | \nexists S_{i,k} \in I_i\ with\ S_{i,k} \succ S_{i,j}\}$$

A series of conditions must be satisfied in order for a compromise to be possible. The first conditions are fairly obvious. As seen in the equation below, the solution we are trying to build a compromise for cannot be Pareto dominant.

$$S_{i,j} \notin P_i$$

Secondly, one of the users, the prospective "winner" of the compromise (henceforth called the consenting user), must approve of the selected solution.

$$u_k \in E_{i,j}$$

Thirdly, the user (henceforth called "the dissenting user") we are trying to compromise with must not approve of the solution. This is somewhat misleading, as the second user is in fact the subject of the search, as it is described below.

$$u_l \in \overline{E}_{i,j}$$

With the two users selected, we must define the condition which makes the compromise possible. Although the solution is not dominant, it would become dominant if the dissenting user would approve. Lacking this condition, the user is not vital to the approval of the solution and is thus not relevant to the search for a compromise.

However, it must be noted that the model can be extended for compromises between more than two users. Indeed, this approach should be pursued if compromises between two users are not available.

$$E'_{i,j} = E_{i,j} \cup \{u_l\} \Rightarrow S_{i,j} \in P_i$$

Since we know our dissenting user's approval would lead to the solution's becoming dominant (in this context we will say that the user is vital for the approval), we must now find something to trade for the approval. As such, we are interested in solutions that the dissenting user has

approved on other issue and which are dependent of the consenting user's approval.
In order to find the set of possible compromises (CS), it is necessary to select all solutions to which the preference of the consenting user is 0 and the preference of the dissenting user is strictly positive.

$$CS_{k,l} = \{S_{m,n} | ipref(u_l, S_{m,n}) > 0, ipref(u_k, S_{m,n}) = 0\}$$

The condition contained in the definition of the CS set is not, however, sufficient. The simple fact that the preferences of the two users differ does not guarantee that the consenting user is vital to those solution's domination. The condition included in the definition of the C (vital compromises) set reflects this by making it necessary that the consenting user's approval would push the solution into the dominant solutions' set. We will say that the compromise is vital for the solution's dominance.

$$C_{k,l} = \{S_{m,n} \in CS_{k,l} | ipref(u_k, S_{i,j}) > ipref(u_l, S_{m,n}), E'_{m,n} = E_{m,n} \cup \{u_k\} \Rightarrow S'_{m,n} \in P_m\}$$

One more condition must be taken into account for the compromise to be viable. Although The set *C* contains possible compromises which are vital to both users, we do not have the guarantee that the compromise is desirable. Hence, the preference of the "winning" user must be larger than that of the "losing" user, in order for compensation to be possible.
While between two users the idea of compensation is simplified, the model can be easily extended to multiple users. The equations remain largely the same, with the exception of the condition for compensation which must be updated the condition that the preference of the winning user must outweigh the sum of the preferences of the losing users.

$$ipref(u_i, S_{i,j}) > \sum_{u \in U'} ipref(u, S_{m,n})$$

The consenting user no longer has to compensate a single user, but multiple users, whose compromises were necessary for the approval of the solution.
Compromises are proposed iteratively through the lifecycle of an issue. Due to this, getting enough approval from the various parties involved in order to actually reach consensus might take a prohibitively large amount of time.
As such, there has to be a way to select those compromises which present the largest chance of being accepted. This can be done by employing a recommendation engine that takes into account the users' prior behavior.

6 Using an action recommendation engine

Since the using a condition such as leading the solution space towards a Paretto optimal state or introducing the idea of compromise through a Kaldor-Hicks model does not guarantee consensus, a mechanism should be introduced which will facilitate grouping of individuals with similar stances. This has a significant chance of reducing the number of solutions through classic social mechanisms (which are outside the system), in the sense that individuals with similar preferences might elect to support each other in order to achieve a greater negotiation strength.
On the other sense, a recommendation system can be used to recommend which solutions, if accepted, would lead to large scale convergence, or in simpler terms, which solutions increase the chance of creating a viable network of compromises.
As a last resort, this would prove useful in applying a traditional democratic mechanism, based on a vote of the majority.
Since recommendation engines are a popular topic due to their large impact on the revenue of internet based businesses, an exhaustive approach of this topic is outside the scope of

this paper. For a much more detailed look at the recommendation engine actually used in the system, see [7] [8].

5 Conclusions

A consensus moderation system would constitute a useful tool in the negotiation between a small number of actors. Consensus in itself serves as a decision mechanism which greatly increases the legitimacy of decisions since all the actors involved have agreed to the final act.

Establishing consensus would be aided by the automatic tool presented in the current paper which would assist participants in aggregating their opinion.

In order to build such a tool, mechanisms borrowed from welfare economics can be employed. The concepts of Pareto domination of proposed solutions and Kaldor Hicks compromises increase the chances of opinion aggregation.

However, since the main interest is to accelerate the process, a recommendation engine which takes into account prior actor behaviour should be employed.

References

[1] P. Speroni di Fenizio, C. Anderson, "Using Pareto Front for a Consensus Building, Human Based, Genetic Algorithm," *Proc. European Conference on Artificial Life (ECAL 2009)*, LNCS, Springer, Berlin, 2009.

[2] http://en.wikipedia.org/wiki/File:Production_Possibilities_Frontier_Curve.svg

[3] http://en.wikipedia.org/wiki/File:Front_pareto.svg

[4] N. Barr, *Economics of the welfare state.* New York, Oxford University Press (USA), 2004.

[5] http://en.wikipedia.org/wiki/File:Kaldor-Hicks-en.svg

[6] http://en.wikipedia.org/wiki/Pareto_efficiency

[7] A. Toma, R. Constantinescu, F. Nastase, "Recommendation system based on the clustering of frequent sets," *WSEAS Transactions On Information Science And Applications*, Volume 6, 2009.

[8] A. Toma, R. Constantinescu, F. Nastase, "Recommendation System Based on the Clustering of Frequent Sets," *8th WSEAS AIKED 09 Proceedings*, 2009.

[9] A. Schotter, *Microeconomics: A Modern Approach*, South-Western College Pub, 2008

[10] D. Collander, *Microeconomics*, McGraw-Hill/Irwin, 2009

Andrei TOMA has a background in both computer science and law and is interested in an interdisciplinary approach to IT Law related issues. He has graduated the Faculty of Cybernetics of the Academy of Economic Studies in Bucharest and the Faculty of Law of the University of Bucharest. He is currently conducting doctoral research at the Academy of Economic Studies.