Conflict Situations Are Inevitable When There Are Many Participants: A Proof Based on the Analysis of Aumann-Shapley Value

Sofia Holguin and Vladik Kreinovich

Abstract When collaboration of several people results in a business success, an important issue is how to fairly divide the gain between the participants. In principle, the solution to this problem is known since the 1950s: natural fairness requirements lead to the so-called Shapley value. However, the computation of Shapley value requires that we can estimate, for each subset of the set of all participants, how much gain they would have gained if they worked together without others. It is possible to perform such estimates when we have a small group of participants, but for a big company with thousands of employers this is not realistic. To deal with such situations, Nobelists Aumann and Shapley came up with a natural continuous approximation to Shapley value – just like a continuous model of a solid body helps, since we cannot take into account all individual atoms. Specifically, they defined the Aumann-Shapley value as a limit of the Shapley value of discrete approximations: in some cases this limit exists, in some it does not. In this paper, we show that, in some reasonable sense, for almost all continuous situations the limit does not exist: we get different values depending on how we refine the discrete approximations. Our conclusion is that in such situations, since computing of fair division is not feasible, conflicts are inevitable.

1 Formulation of the Problem

Collaboration is often beneficial. In many practical tasks, be it menial or intellectual tasks, it is beneficial for several people to collaborate. This way, every participant is focusing on the task in which he/she is most skilled while tasks at which this participant is not very skilled are performed by those who are better in these tasks. In such situations, in general, the more people participate, the better the result.

Sofia Holguin and Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, e-mail: seholguin2@miners.utep.edu, vladik@utep.edu

Question: how to divide the resulting gain. Often, the resulting gain is financial: the company gets a profit, a research group gets a bonus or an award, etc. A natural question is: what is the fair way to divide this gain between the participants.

Shapley value: a description of a fair division. There is a known answer to this question, the answer originally produced by the Nobelist Lloyd Shapley [3, 4, 5]; let us describe this answer.

Let n denote the number of collaborators, and let us number them by numbers from 1 to n. To describe the contribution of each participant, for each subset $S \subseteq N \stackrel{\text{def}}{=} \{1,\ldots,n\}$, we can estimate the gain $v(S)$ that participants from the set S would have gained if they worked on this project without any help from the others. Based on the function $v(S)$ – that assigns, to each subset S, the gain value $v(S)$ – we need to determine how to divide the overall gain $v(N)$ between the participants, i.e., how to come up with the values $x_i(v)$ for which

$$x_1(v)+\ldots+x_n(v) = v(N).$$

Shapley introduced natural requirements on the function $x_i(v)$. First, this function should not depend on the numbers that we assign to the participants: if we start with a different participant etc., each participant should receive the same portion as before. Second, people may participate in two different collaborative projects, corresponding to functions $u(S)$ and $v(S)$. As a result of the first project, each participant i gets $x_i(u)$; as the result of the second project, this participant gets $x_i(v)$. Thus, the overall amount gained by the i-th participant is $x_i(u)+x_i(v)$. Alternatively, we can view the two projects as two parts of one big project. In this case, for each set S, the gain $w(S)$ is equal to the sum of their gains in the two parts: $w(S) = u(S)+v(S)$. Based on this overall project, the i-th participant should get the value $x_i(w) = x_i(u+v)$. It is reasonable to require that the portion assign to the i-th participant should not depend on whether we treat two projects separately or as two parts of a big project. In other words, we should have $x_i(u+v) = x_i(u)+x_i(v)$. This property is known as *additivity*.

Shapley has shown that these two natural requirements uniquely determine the function $x_i(v)$: namely, we should have

$$x_i(v) = \sum_{S:\ i\notin S} \frac{|S|!\cdot(n-|S|-1)!}{n!}\cdot(v(S\cup\{i\})-v(S)),$$

where $|S|$ denotes the number of elements in the set S. This formula is known as *Shapley value*.

Continuous approximation: idea. As we have mentioned, for large number of participants, using the above formula to compute the Shapley value is not feasible. Such situations – when a large number of objects makes computations difficult – are common in science. For example, we know that a solid body consists of molecules, and we know the equations that describe how the molecules interact. However, for a large number of molecules forming a body it is not realistic to take all the molecules into account. Instead, we use a continuous approximation: we assume that the matter is uniformly distributed, and perform computations based on this assumption.

Continuous approximation: towards a precise formulation of the problem. In the continuous approximation, the set of participants forms an area A in 1-D or 2-D or multi-dimensional space. For example, if we want to consider the best way to divide the budget surplus in the state of Texas between projects benefiting local communities, we can view the set of participation as the 2-D area of the state. If in the same task, we want to treat people with different income level differently, it make sense to consider the set of participants as a 3-D region, in which the first two parameters are geographic coordinates, and the third parameter is the income. If we take more parameters into account, we get an area of larger dimension.

As we mentioned earlier, to describe the situation, we assign, to each subset S of the original set of participants, the value $v(S)$ describing how much participants from this set can gain if they act together without using any collaboration from others. In the continuous approximation, we consider reasonable (e.g., measurable) subsets S of the original area A. We also consider reasonable functions $v(S)$, e.g., functions of the type $F(v_1(S),\ldots,v_k(S))$, where $F(z_1,\ldots,z_k)$ is a continuous function, and each v_i has the form $v_i(S) = \int_S f_i(x)\,dx$ for some bounded measurable function $f_i(x)$. Let us denote the set of all such functions by V.

This set V is closed under addition and under a multiplication by a positive number, i.e., if $u,v \in V$ and $\alpha > 0$, then $u+v \in V$ and $\alpha \cdot v \in V$.

Continuous approximation: towards a natural solution. In the original formulation, based on the known values $v(S)$ corresponding to different sets $S \subseteq N$, we decide what portion $x_i(v)$ of the gain $v(N)$ to allocated to each participant i. Once we decide on this, to each group $S \subseteq N$, we thus allocate the sum $x_S(v)$ of values x_i allocated to all the members of this group:

$$x_S(v) = \sum_{i \in S} x_i(v).$$

How can we extend this formula to the continuous case? To do this, we can use the experience of physicists who use the continuous approximation to predict the properties of a solid body. Their technique – see, e.g., [2, 6] – is to discretize the space, i.e., to divide the area occupied by the solid body into small cells, to approximate the behavior of each physical quantity inside a cell by a few parameters (e.g., we assume that this value is constant throughout this small cell), and to solve the corresponding finite-parametric problem. In other words, the usual idea is to provide a discrete approximation to the continuous approximation.

The Nobelists Aumann and Shapley proposed, in effect, to apply the same idea to the continuous approximation to the gain-dividing problem [1]. To approximate the continuous game with a sequence of discrete games, they proposed to consider sequences of partitions $P^{(1)},P^{(2)},\ldots,P^{(k)},\ldots$ each of which divides the area A into a finite number of disjoint measurable sub-areas $A^{(k)}_1,\ldots,A^{(k)}_{n_k}$, and that satisfy the following two properties:

- the next division $P^{(k+1)}$ is obtained from the previous division $P^{(k)}$ by subdividing each of the sets $A_i^{(k)}$, and
- for every two elements $a \neq b$ from the area A, there exists a number k for which the partition $P^{(k)}$ allocates these two elements to different sub-areas.

Such sequences of partitions are called *admissible*.

For each partition $P^{(k)}$ for an admissible sequence, we consider sets S consisting of the corresponding sub-areas, i.e., sets S of the type

$$S(s) = \bigcup_{i \in s} A_i^{(k)}$$

for some set $s \subseteq \{1, \ldots, n_k\}$. In this discrete-approximation-to-continuous-approximation scheme, we get, for each k, a situation with n_k participants for which the gain of each subset $s \subseteq \{1, \ldots, n_k\}$ is described by the value $v(S(s))$. Based on these values, we can compute, for each of these participants, the Shapley value $x_i^{(k)}(v)$.

For each measurable subset $S \subseteq A$ and for each partition $P^{(k)}$, we can thus find the approximate lower bound for the amount allocated to S as the sum

$$x_S^{(k)}(v) = \sum_{i:\, A_i^{(k)} \subseteq S} x_i^{(k)}(v).$$

Some functions $v(S)$ have the following nice property: for each set S, no matter what admissible sequence of partitions $P^{(k)}$ we take, the values $x_S^{(k)}(v)$ tend the same limit. This limit $x_S(v)$ is called the *Aumann-Shapley value* corresponding to the function $v(S)$.

Sometimes, the Aumann-Shapley value exists, sometimes, it does not exist. For some functions $v(S)$, the sequences $x_S^{(k)}(v)$ always converge. So, if the original situation with a large number of participants can be described by such function $v(S)$, then, to fairly allocate gains to all the participants, we can use an appropriate simplified approximation for which computing such a fair allocation is feasible.

On the other hand, it is known that for some functions $v(S)$, the sequences $x_S^{(k)}(v)$ do not converge. For example, Example 19.2 from [1] shows that the limits do not exist already for the following simple function on the domain $A = [0, 1]$:

$$v(S) = \min\left(\int_S \chi_{[0,\frac{1}{3})}\,dx, \int_S \chi_{[\frac{1}{3},\frac{2}{3})}\,dx, \int_S \chi_{[\frac{2}{3},1]}\,dx\right),$$

where χ_I is the characteristic function of the set I, i.e., $\chi_I(x) = 1$ if $x \in I$ and $\chi_I(x) = 0$ is $x \notin I$. For such functions v, we cannot use a simplified approximation.

And since for large number of participants, direct computation of the Shapley value – i.e., of the fair division – is not feasible, this means that for such functions $v(S)$, we cannot feasibly compute a division that everyone would recognize as fair. Thus, in such situations, conflicts are inevitable.

Natural question. A natural question is: which of these two situations occurs in real life?

What we do in this paper. In this paper, we come up with an answer to this question.

2 Discussion and the Resulting Conclusion

Main idea behind our analysis. To get an answer to the above question, we will use the following principle that physicists use (see, e.g., [2, 6]): that if some phenomenon occurs in almost all situations ("almost all" in some reasonable sense), then they conclude that this phenomenon occurs in real life as well.

For example, if you flip a coin many times and select 1 when it falls head and 0 when it falls tail, in principle, you can get any sequence of 0s and 1s. However, we know that for almost all sequences of 0s and 1s, the frequency of 1s tends to 1/2. So, physicists conclude (and experiments confirm this) that when we flip a coin many times, the frequency of 1s tends to 1/2. This is not just what physicists do, this is common sense: if you go to a casino and the roulette ends up on red (as opposed to black) 30 times in a row, you will naturally conclude that the roulette is biased.

Similarly: it is, in principle, possible that due to random thermal interactions between all the molecules in a cat's body, all the molecules will start moving up, and the poor cat will start rising in the air. However, the probability of this event is practically 0. In almost all cases, this is not possible, so physicists conclude that this is not possible in real life.

An even simpler example: if we run some stochastic process twice and we get the exact same result in both cases, this would mean that something is wrong: indeed, once we have the first result r_1, the second result r_2 can, in principle, take any real value. Only for one of these values r_1, out of continuum many, we can have $r_2 = r_1$. Thus, for almost all possible values r_2 (with one exception), we have $r_2 \neq r_1$; thus, we conclude that in real life, we will have $r_2 \neq r_1$.

In our analysis, we will take into account additivity and homogeneity of the Aumann-Shapley value. It is known that several properties of the Shapley value can be extended to the Aumann-Shapley value: we just need to take into account that, in contrast to the Shapley value – which is always defined – the Aumann-Shapley value is only defined for some functions $v(S)$. As we have mentioned, the Shapley value has the additivity property: for every two functions $u(S)$ and $v(S)$, we have $x_i(u+v) = x_i(u) + x_i(v)$. Also, it has the following homogeneity property: for every $\alpha > 0$, we have $x_i(\alpha \cdot v) = \alpha \cdot x_i(v)$. In the limit, these two properties leads to the following properties of the Aumann-Shapley (AS) value:

- if the AS value exists for functions u and v, then AS value exists for $u+v$, and

$$x_S(u+v) = x_S(u) + x_S(v);$$

- if the AS value exists for functions u and $u+v$, then AS value exists for v, and

$$x_S(u+v) = x_S(u) + x_S(v);$$

- if the AS value exists for v, then for each $\alpha > 0$, the AS value exists for $\alpha \cdot v$, and

$$x_S(\alpha \cdot v) = \alpha \cdot x_S(v).$$

What we can conclude from these properties. Let v_0 be a function for which the AS value does not exist. For each function $v \in V$, we can consider a 1-D set $L(v) \stackrel{\text{def}}{=} \{v + t \cdot v_0 \,|\, v + t \cdot v_0 \in V\}$, where t denotes an arbitrary real number. For all $t \geq 0$, we have $v + t \cdot v_0 \in V$, so this set contains a whole half-line, i.e., it contains continuum many points. Let us prove that in this set, we can have at most one real value t for which the AS value exist.

The proof is by contradiction. Indeed, if we have two values $t_1 < t_2$ for which the functions $v + t_i \cdot v_0$ have AS value, then, by additivity, the AS value exists for the difference $(v + t_2 \cdot v_0) - (v + t_1 \cdot v_0) = (t_2 - t_1) \cdot v_0$ and thus, by homogeneity, for the function $(t_2 - t_1)^{-1} \cdot ((t_2 - t_1) \cdot v_0) = v_0$. However, we have selected v_0 for which the AS value does not exist. The resulting contradiction proves our statement.

This means that for almost all functions v, the AS value does not exist. One can easily check that for two functions v and v', if the corresponding sets $L(v)$ and $L(v')$ have a common element, then these sets coincide. Thus, we can divide the whole set V of all possible functions $v(S)$ into non-intersecting 1-D subsets $L(v)$ corresponding to different v. In each such subset, out of continuum many points, there is at most one point for which the AS value exists.

When some property holds only for one point on the whole half-line, this means, intuitively, that in almost all cases, this property is not satisfied. This is exactly what we wanted to prove.

Corollary: reminder. We have shown that for almost all functions $v(S)$, the AS value does not exist – i.e., that the Shapley values corresponding to discrete few-player approximations do not converge. Following the above physicists' principle, we conclude that in real life, these values do not converge. Thus, in situations with many participants, we cannot use this approximation idea.

And since for large number of participants, direct computation of the Shapley value (i.e., of the fair division) is not feasible, this means that for such functions $v(S)$, we cannot feasibly compute a division that everyone would recognize as fair. Thus, in such situations, conflicts are inevitable.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in

Computer Science), HRD-1834620 and HRD-2034030 (CAHSI Includes), EAR-2225395, and by the AT&T Fellowship in Information Technology.

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

References

1. R. J. Aumann and L. S. Shapley, *Values of Non-Atomic Games*, Princeton University Press, Princeton, New Jersey, 1974.
2. R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.
3. M. Maschler, E. Solan, and S. Zamir, *Game Theory*, Cambridge University Press, Cambridge, UK, 2020.
4. L. S. Shapley, *Notes on the n-Person Game – II: The Value of an n-Person Game*, RAND Corporation Research Memorandum RM-670, Santa Monica, California, 1951, https://www.rand.org/content/dam/rand/pubs/research_memoranda/2008/RM670.pdf
5. L. S. Shapley, "A value of *n*-person games", In: H. W. Kuhn and A. W. Tucker (eds.), *Contributions to the Theory of Games*, Princeton University Press, Princeton, New Jersey, 1953, pp. 307–317.
6. K. S. Thorne and R. D. Blandford, *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics*, Princeton University Press, Princeton, New Jersey, 2021.