Fair Economic Division: How to Modify Shapley Value to Take Into Account that Different People Have Different Productivity

Abstract

Purpose: When several participants, working together, gained some amount of money, what is the fair way to distribute this amount between them? This is the problem that the future Nobelist Lloyd Shapley was working on when he proposed what is now called the Shapley value – a division uniquely determined by natural fairness assumptions. However, this solutions is not universal: it assumes that all participants are equal – in particular, that they have equal productivity. In practice, people have different productivity levels, and these productivity levels can differ a lot: e.g., some software engineers are several times more productive than others. It is desirable to take this difference in productivity into account.

Design/methodology/approach: Shapley value is based on an axiomatic approach: it is uniquely determined by the appropriate fairness assumptions. To generalize Shapley value to the case of different productivity, we modified these assumptions appropriately, and analyzed what can be derived from these modified assumptions.

Findings: We prove that there is a unique division scheme that satisfies all the resulting assumptions. This scheme is thus a generalization of Shapley value to this more general and more realistic situation, when different participants have different productivity.

Originality/value: Both the formulation of the problem and the result are new. The resulting division scheme can be used to more adequately distribute the common gains – by explicitly taking into account that different participants have, in general, different productivity.

Keywords: Shapley value, Fair distribution, Difference in productivity, Axiomatic approach.

1 Formulation of the problem

A simple motivating example. Suppose that three people came up together with a code that they sold to a major software company for 1 million US dollars. They worked all together, so without all three of them, this unfinished code would be worth nothing. Suppose also that the productivity of Person 1 is twice higher than the productivity of each of his two colleagues. In this situation, what is a fair way to distribute the money between these three persons?

Fair division: a problem. The above simple example is a particular case of the general problem of fair division. Let us describe this general problem in precise terms.

Let us assume that a group $N_n \stackrel{\text{def}}{=} \{1, \ldots, n\}$ of n people jointly gets some benefit $v(N_n)$. What is the fair way to distribute this benefit between the participants, i.e., to assign values $\varphi_1, \ldots, \varphi_n$ whose sum is $v(N_n)$?

What do we need to know to make a fair division. To make a fair distribution, it is important to know what is the contribution of each participant. This can be described by providing, for each subset $S \subseteq N$, a value $v(S)$ that people from S would have gained if they acted on their own, without help of others. So, we get a function $S \mapsto v(S)$ that characterizes the situation.

Shapley value. A Nobelist Lloyd Shapley found out that under reasonable conditions, there is only one way $\varphi_i(v)$ to assign the distribution to each such function $v(S)$; see, e.g., (Shapley, 1951), (Shapley, 1953), (Roth, 1988), (Luce and Raiffa, 1989), (Myerson, 1997), and (Owen, 2013). This distribution has many equivalent forms, In this paper, we use the following form

$$\varphi_i(v) = \sum_{S: i \in S} \frac{t(S)}{|S|}, \tag{1}$$

where $|S|$ denotes the number of elements in the set S and

$$t(S) \stackrel{\text{def}}{=} \sum_{R \subseteq S} (-1)^{|S|-|R|} \cdot v(R). \tag{2}$$

Comment. It is worth mentioning that the transformation (2) is well-known:

- in random set theory (see, e.g., (Molchanov, 2017) and (Nguyen, 2006)), where it is used to transform the cumulative distribution function into the corresponding probability density function, and
- in Dempster-Shafer theory (see, e.g., (Shafer 1976), (Nguyen, 2006)), where it is used to transform the belief function into the corresponding mass function.

What are the requirements behind Shapley value. Shapley's first condition is *symmetry*: if two participants i and j contribute equally, i.e., if the values $v(S)$ do not change when we swap i and j, then these participants should get equal amounts: $\varphi_i(v) = \varphi_j(v)$.

Shapley's second condition is that if a person i is not contributing, i.e., if $v(S \cup \{i\}) = v(S)$ for all S, then we should have $\varphi_i(v) = 0$.

Shapley's third condition is additivity: if we have two situations $u(S)$ and $v(S)$, then we can:

- either consider them separately
- or view them as a single situation with gain $w(S) = u(S) + v(S)$.

The outcome should not depend on how we view this, so we should have $\varphi_i(w) = \varphi_i(u) + \varphi_i(v)$.

Comment. In this paper, we will only deal with economic applications of Shapley value. It should be mentioned that Shapley value is now also actively used in machine learning, to find the importance $\varphi_i(v)$ of each of n features based on the effectiveness $v(S)$ of solving the problem when we only use features from the set S.

Why go beyond Shapley value. At first glance, the division provided by the Shapley value sounds reasonable – and this division has been indeed successfully applied to many practical problems. We will show, however, that this division is not always reasonable and fair. To show this, let us go back to the simple example with which we started this paper. In this example, the three persons worked all together, so without all three of them, this unfinished code would be worth nothing. Thus, here $v(\{1,2,3\}) = 1$ while $v(S) = 0$ for all proper subsets of the set $\{1,2,3\}$.

Let us analyze what happens if we apply Shapley value to this situation. In this case, $t(\{1,2,3\}_{=}1$ and $t(S) = 0$ for all other subsets $S \subseteq \{1,2,3\}$. Thus, the Shapley value formula leads to

$$\varphi_1 = \varphi_2 = \varphi_3 = \frac{t(\{1,2,3\})}{3} = \frac{1}{3}.$$

So, each of the three persons would receive exactly one third of a million – which is not fair, since Person 1 did twice more work than Person 2 or Person 3.

The reason why we got this unfair division is that one of the requirements on which Shapley's value is based is the requirement of symmetry. In this example, the function $v(S)$ is invariant with respect to all possible permutations – and, as a result, the division is also invariant under each permutation, i.e., all three persons get the same amount.

So, this unfairness was caused by the fact that in the Shapley value setting, we only take into account the function $v(S)$, and we do not take into account the additional information – that these three persons have different productivity. It is therefore desirable to take productivity p_i of each participant into account. Productivity can be measured in the usual way – by the amount $p_i > 0$ of the work that Person i can perform during a given time duration – e.g., how many lines of code the person can produce during an 8-hour working day. In other words, we need to determine the values φ_i

based on both $v(S)$ and the values $p = (p_1, p_2, \ldots)$: $\varphi_i(v, p)$.

Comment. In game theory, descriptions in which we only take into account the function $v(S)$ are known as *coalitional games*. What we are suggesting is to add additional information to the description of the situation. Situations when we supplement the function $v(S)$ with additional information are known as *coalitional games with prior additional information.*

What we do in this paper. In this paper, we show how to adjust the requirement behind the Shapley value so that they would lead to a unique determination of the desired distributions $\varphi_i(v, p)$.

2 What we propose

Natural first requirement. If i's productivity is twice larger than j's, this means that the company can replace i with two less productive workers and get the same result. After this replacement, all participants have the same productivity, so to this replaced situation, we can apply symmetry and get Shapley value – and then assign to i the sum of bonuses that Shapley value recommends for his/her two replacements.

Similarly, if we could replace person i with 3 or more workers and keep the same productivity, it makes sense to require that the amount given to the person i should be equal to the sum of amounts given to these workers. It turns out that it is sufficient to require this property only for situation which are *simple* – in some precise sense described below.

Natural second requirement. If the productivity changes a little bit, the resulting distribution should also change only a little bit. In mathematical terms, this means that the dependence of distribution on productivity should be continuous.

Main result. If impose these two additional requirements, then we get the following result.

Definition 1. *By a* situation*, we mean a triple* (n, v, p)*, where:*

- n *is a positive integer,*

- *v is a function that assigns a value $v(S) \geq 0$ to each subset $S \subseteq N_n \stackrel{\text{def}}{=} \{1, \ldots, n\}$, and for which $R \subseteq S$ implies $v(R) \leq v(S)$, and*

- *$p = (p_1, \ldots, p_n)$ is a tuple of n positive numbers.*

Definition 2. *We say that the situation (n, v, p) is* simple *if for some subset $R \subseteq N_n$, we have $v(S) = 0$ if $R \not\subseteq S$ and $v(S) = v(R)$ otherwise. We will call this set R* basic.

Definition 3.

- *By a* division strategy*, we mean a function $\varphi(n, v, p)$ that assigns, to each situation (n, v, p), an n-tuple of real numbers $\varphi_i(n, v, p)$, $1 \leq i \leq n$, for which $\varphi_1(n, v, p) + \ldots + \varphi_n(n, v, p) = v(N_n)$.*

- *We say that a division strategy is* symmetric *if for every situation in which swapping i and j does not change v and p, i.e., in which $p_i = p_j$ and $v(S \cup \{i\}) = v(S \cup \{j\})$ for all sets S that contain neither i nor j, we have $\varphi_i(n, v, p) = \varphi_j(n, v, p)$.*

- *We say that a division strategy has the* null-player property *if for each situation and for each player i for which $v(S \cup \{i\}) = v(S)$ for all S, we get $\varphi_i(n, v, p) = 0$.*

- *We say that a division strategy is* additive *if for all u and v, we have $\varphi_i(n, u + v, p) = \varphi_i(n, u, p) + \varphi_i(n, v, p)$ for all i.*

- *We say that a division strategy is* continuous *if $\varphi(n, v, p)$ ia a continuous function of p.*

- *We say that a division strategy is* productivity-based *if for every simple situation with a basic set R, if we combine participants from a subset $R' \subseteq R$ into a single participant i_0 with productivity equal to the sum of productivities of all members of R', then in this new situation (n', v', p'),*

$$\varphi_{i_0}(n', v', p') = \sum_{i \in R'} \varphi_i(n, v, p).$$

Proposition. *There exists one and only one division strategy which is symmetric, has null-player property, is additive, continuous, and productivity-based. In this strategy,*

$$\varphi(n, v, p) = p_i \cdot \sum_{S: i \in S} \frac{t(S)}{p(S)}, \tag{3}$$

where we denoted

$$p(S) \stackrel{\text{def}}{=} \sum_{i \in S} p_i. \tag{4}$$

Comments.

- The proof of this proposition is given in the next section.
- One can easily see that if all participants have the same productivity, i.e., if $p_1 = \ldots = p_n$, then the formula (3) becomes the usual Shapley value formula (1).

Let us apply this result to our simple example. Let us see what happens if we apply this result to our simple example, where three persons – one of which has twice higher productivity than each of the two others – need to divide one million dollars that they got from the major software company for their code. If we simply apply Shapley value to this situation, then each of them would receive exactly 1/3 of a million, which is not fair since Person 1 did twice more work than Person 2.

Let us analyze what happens if we apply the newly proposed scheme – the scheme that takes difference in productivity into account – to this example. Here, if we take the productivity of Persons 2 and 3 as the measuring unit, i.e., take $p_2 = p_3 = 1$, then the productivity of Person 1 is twice larger: $p_1 = 2$. Here, $p(S) = p_1 + p_2 + p_3 = 2 + 1 + 1 = 4$. So, according to the formula (3), we have

$$\varphi_1 = p_1 \cdot \frac{t(\{1, 2, 3\})}{p(\{1, 2, 3\})} = 2 \cdot \frac{1}{4} = 0.5,$$

$$\varphi_2 = p_2 \cdot \frac{t(\{1,2,3\})}{p(\{1,2,3\})} = 1 \cdot \frac{1}{4} = 0.25, \text{ and}$$

$$\varphi_3 = p_3 \cdot \frac{t(\{1,2,3\})}{p(\{1,2,3\})} = 1 \cdot \frac{1}{4} = 0.25.$$

So, in the new scheme, Person 1 will get twice more than each of 2 and 3: Person 1 will get \$500,000, while Persons 2 and 3 will get \$250,000 each.

A slightly more complex example. The above example is simple, the resulting fair division can be proposed without any Shapley-like formulas. The main purpose of this simple example was to show that there is a problem with the naive application of Shapley value — since the Shapley value does not take into account difference in productivity.

Our new method allow to deal with more complex situations, when we have both non-zero values of $v(S)$ for different sets S, and information about productivity. Let us illustrate this method on a slightly more complex example.

Let us assume that in addition to Persons 1, 2, and 3 designing a software that is worth \$1 million, a somewhat different group – Persons 1, 2, and 4 – developed a related software that is also worth \$1 million. Here, Person 1 still has twice larger productivity than Persons 2 and 3, and Person 4 has the same productivity as Persons 2 and 3. These two softwares supplement each other, so when presented together they are worth more – when presented as a package, they are worth \$2.5 million. What is then a fair way to divide \$2.5 million between the four persons?

In this case, the function $v(S)$ has the following form: $v(\{1,2,3\}) = v(\{1,2,4\}) = 1$, $v(\{1,2,3,4\}) = 2.5$, and $v(S) = 0$ for all other sets $S \subseteq \{1,2,3,4\}$. One can check that in this case, we have $t(\{1,2,3\}) = t(\{1,2,4\}) = 1$, $t(\{1,2,3,4\}) = 0.5$, and $t(S) = 0$ for all other sets $S \subseteq \{1,2,3,4\}$.

Thus, in the division based on the Shapley value, each of the Persons 1 and 2 gets

$$\varphi_1 = \varphi_2 = \frac{t(\{1,2,3\})}{3} + \frac{t(\{1,2,4\})}{3} + \frac{t(\{1,2,3,4\})}{4} =$$

$$\frac{1}{3} + \frac{1}{3} + \frac{0.5}{4} = \frac{2}{3} + \frac{1}{8} = \frac{19}{24} \approx 0.79,$$

while each of the Persons 3 and 4 gets:

$$\varphi_3 = \frac{t(\{1,2,3\})}{3} + \frac{t(\{1,2,3,4\})}{4} = \frac{1}{3} + \frac{0.5}{4} = \frac{1}{3} + \frac{1}{8} = \frac{11}{24} \approx 0.46 \text{ and}$$

$$\varphi_4 = \frac{t(\{1,2,4\})}{3} + \frac{t(\{1,2,3,4\})}{4} = \frac{1}{3} + \frac{0.5}{4} = \frac{1}{3} + \frac{1}{8} = \frac{11}{24} \approx 0.46.$$

This is not fair, since Person 1 worked twice harder than Person 2, but they get the same amount.

In the new approach, with $p_1 = 2$ and $p_2 = p_3 = p_4$, we have $p(\{1,2,3\}) = p_1 + p_2 + p_3 = 2 + 1 + 1 = 4$, $p(\{1,2,4\}) = p_1 + p_2 + p_4 = 2 + 1 + 1 = 4$, and $p(\{1,2,3,4\}) = p_1 + p_2 + p_3 + p_4 = 2 + 1 + 1 + 1 = 5$. Thus, we have

$$\varphi_1 = p_1 \cdot \left(\frac{t(\{1,2,3\})}{p(\{1,2,3\})} + \frac{t(\{1,2,4\})}{p(\{1,2,4\})} + \frac{t(\{1,2,3,4\})}{p(\{1,2,3,4\})} \right) =$$

$$2 \cdot \left(\frac{1}{4} + \frac{1}{4} + \frac{0.5}{5} \right) = 2 \cdot (0.25 + 0.25 + 0.1) = 1.2,$$

$$\varphi_2 = p_2 \cdot \left(\frac{t(\{1,2,3\})}{p(\{1,2,3\})} + \frac{t(\{1,2,4\})}{p(\{1,2,4\})} + \frac{t(\{1,2,3,4\})}{p(\{1,2,3,4\})} \right) =$$

$$1 \cdot \left(\frac{1}{4} + \frac{1}{4} + \frac{0.5}{5} \right) = 1 \cdot (0.25 + 0.25 + 0.1) = 0.6,$$

$$\varphi_3 = p_3 \cdot \left(\frac{t(\{1,2,3\})}{p(\{1,2,3\})} + \frac{t(\{1,2,3,4\})}{p(\{1,2,3,4\})} \right) =$$

$$1 \cdot \left(\frac{1}{4} + \frac{0.5}{5} \right) = 1 \cdot (0.25 + 0.1) = 0.35, \text{ and}$$

$$\varphi_4 = p_4 \cdot \left(\frac{t(\{1,2,4\})}{p(\{1,2,4\})} + \frac{t(\{1,2,3,4\})}{p(\{1,2,3,4\})} \right) =$$

$$1 \cdot \left(\frac{1}{4} + \frac{0.5}{5} \right) = 1 \cdot (0.25 + 0.1) = 0.35.$$

Here, Person 1 gets twice more than Person 2 – in agreement with the fact that Person 1 did twice as much work as Person 2.

3 Proof

1°. It is relatively easy to prove that the strategy (3) satisfied all the above properties. So, to complete the proof, it is sufficient to prove that if a division strategy satisfies all the above properties, then it has the form (3).

2°. Let us assume that a division strategy satisfies all the above properties. Let us first show that for each simple situation (n, v, p) with a basic set R:

- for all $i \notin R$, we have $\varphi_i(n, v, p) = 0$, and
- for all $i \in R$, we have

$$\varphi_i(n, v, p) = p_i \cdot \frac{v(R)}{p(R)}. \quad (5)$$

Let us prove these two formulas one by one.

2.1°. From the null-player property, it follows that $\varphi_i(n, v, p) = 0$ for all $i \notin R$ – since these participants do not add anything to any value $v(S)$.

2.2°. Let us first prove the formula (5) for the cases when all the productivity values are rational numbers, i.e., ratios of natural numbers. In this case, we can bring all these rational numbers to a common denomination d, so we have

$$p_i = \frac{n_i}{d}$$

for some integers n_i. In this case, we can replace each participant i with n_i participants with productivity $1/d$. After this substitution, we get a new situation with $\sum n_j$ participants with the same productivity. So, due to symmetry, each of them will get the exact same share, i.e., the value

$$\frac{v(N)}{\sum\limits_j n_j} = \frac{v(R)}{\sum\limits_j n_j}.$$

Due to the fact the division is productivity-based, each participants i in the original situation gets n_i times more, i.e., gets the value

$$\varphi(n,v,p)=n_i\cdot\frac{v(R)}{\sum\limits_j n_j}.$$

We can divide both the numerator and the denominator of this expression by d, this will not change the resulting values. So, we get

$$\varphi(n,v,p)=\frac{n_i}{d}\cdot\frac{v(R)}{\sum\limits_j \frac{n_j}{d}},$$

i.e.,

$$\varphi(n,v,p)=p_i\cdot\frac{v(R)}{\sum\limits_j p_j}.$$

By definition of $p(R)$, this means that

$$\varphi(n,v,p)=p_i\cdot\frac{v(R)}{p(R)}.$$

We have proved the formula (5) for rational productivity values p_i. Any real-valued productivities can be approximated, with any given accuracy, by rational numbers. Thus, in the limit when these rational approximations tend to the original values p_i, we conclude – dut to continuity – that the formula (5) holds for real values as well.

3°. It is known that for every function v, we have

$$v(S)=\sum_{R\subseteq S} t(R).$$

Some of the value $t(R)$ may be negative. If we move these values to the left-hand side, we get the following equality

$$v(S)+\sum_{R\subseteq S:t(R)<0}|t(R)|=\sum_{R\subseteq S:t(R)\ge 0} t(R).$$

For every set R, let us consider a simple situation v_R with this basic set R

and $v_R(R) = |t(R)|$, then we have

$$v(S) + \sum_{R \subseteq S: t(R) < 0} v_R(S) = \sum_{R \subseteq S: t(R) \geq 0} v_R(S).$$

Thus, by additivity,

$$\varphi_i(n, v, p) + \sum_{R: i \in R \,\&\, t(R) < 0} \varphi_i(n, v_R, p) = \sum_{R: i \in R \,\&\, t(R) \geq 0} \varphi_i(n, v_R, p).$$

From Part 2 of this proof, we know the values $\varphi(n, v_R, p)$ for all simple situations. Thus, we get

$$\varphi_i(n, v, p) + \sum_{R: i \in R \,\&\, t(R) < 0} p_i \cdot \frac{|t(R)|}{p(R)} = \sum_{R: i \in R \,\&\, t(R) \geq 0} p_i \cdot \frac{t(R)}{p(R)}.$$

If we move the sum from the left-hand side into the right-hand side, we get the formula

$$\varphi_i(n, v, p) = \sum_{R: i \in R} p_i \cdot \frac{t(R)}{p(R)}.$$

If we have the common factor out of the sum, we get the desired formula (3).

The proposition is proven.

4 Conclusions

In many practical situations, after a group of people has jointly performed an important task, it is necessary to fairly distribute the resulting benefits between them. Such situations were motivations for Nobelist Lloyd Shapley, who showed that under some reasonable conditions, there is only one fair way to distribute these benefits – which we now call Shapley value. Shapley value is based on the available information about contribution of subgroups of participants – gauged by how much each subgroup could gain if it acted on its own, without help of others.

In this paper, we show that while Shapley value has been successfully applied to many economic and financial situations, there are cases when it does not lead to a fair division. The reason for this is that Shapley value does not explicitly take into account that participants often have different productivity. In some cases, this difference is taken into account implicitly – since different productivity levels affect the subgroup gains. However, in some cases, subgroup gains are not affected by this difference, and it leads to unfair outcomes – when more productive participants get paid the same amount as less productive ones. To avoid such situations, we show how to take productivities into account. Specifically, we show that, in this case, a natural generalization of Shapley's conditions also leads to a unique way to distribute benefits. The resulting formulas are illustrated on two simple realistic examples.

Acknowledgments

The authors are thankful to Nguyen Ngoc Thach for his encouragement and advice, and to the anonymous referees for valuable suggestions.

References

R. D. Luce and R. Raiffa, *Games and Decisions: Introduction and Critical Survey*, Dover, New York, 1989.

I. Molchanov, *Theory of Random Sets*, Springer, Cham, Switzerland, 2017.

R. B. Myerson, *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, Massachusetts, 1997.

H. T. Nguyen, *An Introduction to Random Sets*, Chapman & Hall/CRC, Boca Raton, Florida, 2006.

G. Owen, *Game Theory*, Emerald Publishers, Bingley, UK, 2013.

A. E. Roth (ed.), *The Shapley Value: Essays in Honor of Lloyd S. Shapley*, Cambridge University Press, Cambridge, 1988. doi:10.1017/CBO9780511528446, ISBN 0-521-36177-X.

G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, New Jersey, 1976.

L. S. Shapley, *Notes on the n-Person Game – II: The Value of an n-Person Game*, Research Memorandum RM-670, RAND Corporation, Santa Monica, California, 1951.

L. S. Shapley, "Value for n-person Games", In: H. W. Kuhn and A. W. Tucker (eds.), *Contributions to the Theory of Games*, Annals of Mathematical Studies, Vol. 28, Princeton University Press, Princeton, New Jersey, USA, 1953, pp. 307–317, doi:10.1515/9781400881970-018. ISBN 9781400881970.