5-2020

Reward for Good Performance Works Better Than Punishment for Mistakes: Economic Explanation

Olga Kosheleva

Julio Urenda

Vladik Kreinovich

Comments:

Technical Report: UTEP-CS-20-40

Reward for Good Performance Works Better Than Punishment for Mistakes: Economic Explanation

Olga Kosheleva[1], Julio Urenda[2,3], and Vladik Kreinovich[3]
[1]Department of Teacher Education
[2]Department of Mathematical Sciences
[3]Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
olgak@utep.edu, jcurenda@utep.edu, vladik@utep.edu

Abstract

How should we stimulate people to make them perform better? How should we stimulate students to make them study better? Many experiments have shown that reward for good performance works better than punishment for mistakes. In this paper, we provide a possible theoretical explanation for this empirical fact.

1 Formulation of the Problem

Reward vs. punishment: an important economic problem. One of the most important issue in economics is how to best stimulate people's productivity, what is the best combination of reward and punishment that makes people perform better.

This problem is ubiquitous. This problem rises not only in economics, it appears everywhere. How do we stimulate students to study better? How do we stimulate our own kids to behave better?

Empirical fact. In his famous book [3] summarizing his research, the Nobelist Daniel Kahneman – one of the fathers of behavioral economics – cites numerous researches that all confirm that reward for good performance, in general, works better than punishment for mistakes; see, e.g., p. 175.

But why? Like many other facts from behavioral economics, this empirical fact does not have a convincing theoretical explanation.

What we do in this paper. In this paper, we provide a theoretical explanation for this empirical phenomenon.

2 Analysis of the Problem

What people want. People spend some efforts e, and, based on results of these efforts, they get some reward $r(e)$. In the first approximation, we can say that the overall gain is the reward minus the efforts, i.e., the difference

$$r(e) - e. \tag{1}$$

A natural economic idea is that every person wants to maximize his/her gain, i.e., maximize the difference (1).

How to proceed. In view of the formula (1), to explain why rewards work better than punishments, we need to analyze what are the reward functions $r(e)$ corresponding to these two different reward strategies. Similarly to our derivation of the formula (1), in our analysis, we will use simplified "first approximation" models, i.e., models providing a good qualitative understanding of the situation.

What reward function corresponds to rewarding good performance. Crudely speaking, rewarding good performance means that:

- if the performance is not good, i.e., if the effort e is smaller than the smallest needed effort e_0, there is practically no reward:

$$r(e) = r_+ \tag{1}$$

  for some $r_+ \approx 0$;

- on the other hand, the more effort the person uses, the larger the reward; in other words, every effort beyond e_0 is proportionally rewarded, i.e.,

$$r(e) = r_+ + c_+ \cdot (e - e_0), \tag{2}$$

  for some constant c_+.

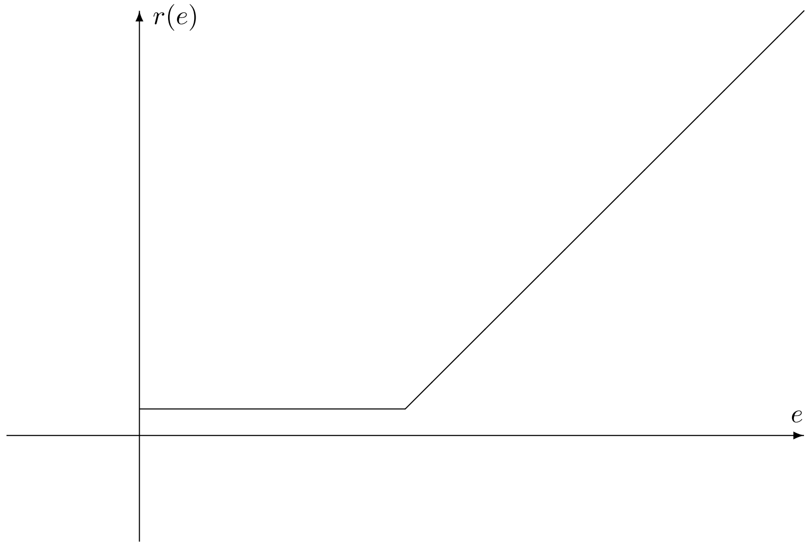
The constant c_+ depends on the units used for measuring effort and reward: one unit of effort corresponds to c_+ units of reward. The two formulas:

- the formula (1) corresponding to the case $e \le e_0$, and
- the formula (2) corresponding to the case $e \ge e_0$,

can be combined into a single formula

$$r(e) = r_+ + \max(0, c_+ \cdot (e - e_0)) = r_+ + c_+ \cdot \max(0, e - e_0). \tag{3}$$

This dependence has the following form:



What can we say about this function. It is easy to see that the function (3) is *convex* in the sense that for all $e' < e''$ and for each $\alpha \in [0, 1]$, we have

$$r(\alpha \cdot e' + (1 - \alpha) \cdot e'') \le \alpha \cdot r(e') + (1 - \alpha) \cdot r(e'').$$

What reward function corresponds to punishing for mistakes. Crudely speaking, punishing for mistakes means that:

- if the performance is good, i.e., if the effort e is larger than or equal to the smallest needed effort e_0, then there is no punishment, i.e., the reward remains the same:
$$r(e) = r_- \quad (4)$$
for some constant r_-;

- on the other hand, the fewer effort the person uses, the most mistakes he/she makes, so the larger the punishment and the smaller the resulting reward; in other words, every effort below e_0 is proportionally penalized, i.e.,
$$r(e) = r_- - c_- \cdot (e_0 - e), \quad (5)$$
for some constant c_-.

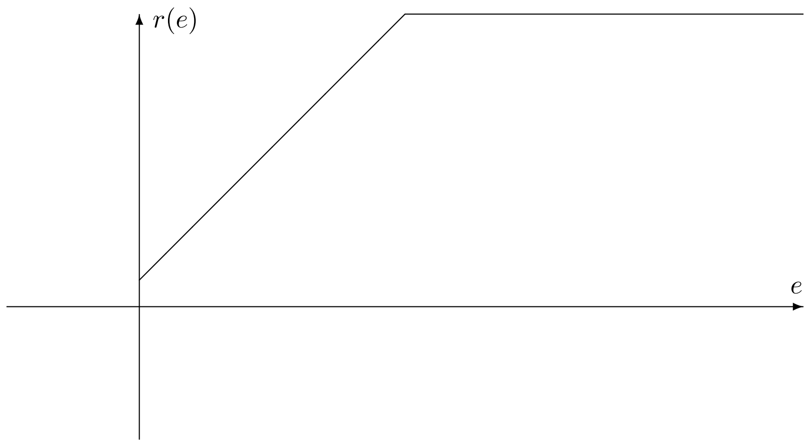
The constant c_- depends on the units used for measuring effort and reward: one unit of effort corresponds to c_- units of reward. The two formulas:

- the formula (4) corresponding to the case $e \ge e_0$, and

- the formula (5) corresponding to the case $e \le e_0$, can be combined into a single formula

$$r(e) = r_- - c_- \cdot \max(0, e_0 - e) = r_- + c_- \cdot \min(0, e - e_0). \tag{6}$$

This dependence has the following form:



What can we say about this function. It is easy to see that this function is *concave* in the sense that for all $E' < E''$ and for each $\alpha \in [0, 1]$, we have

$$r(\alpha \cdot e' + (1 - \alpha) \cdot e'') \geq \alpha \cdot r(e') + (1 - \alpha) \cdot r(e'').$$

Now, we are ready to present the desired explanation.

3 Our Explanation

Known properties of convex and concave functions: reminder. It is known (see, e.g., [5]) that:

- every linear function is both convex and concave;
- the sum of two convex functions is convex, and
- the sum of two concave functions is concave.

In particular, the linear function $f(e) = -e$ is both convex and concave, thus:

- when the function $r(e)$ is convex, the sum $r(e) + (-e) = r(e) - e$ is also convex; and
- when the function $r(e)$ is concave, the sum $r(e) + (-e) = r(e) - e$ is also concave.

It is also known that:

- for a convex function, the maximum on an interval is always attained at one of this interval's endpoints, while
- for a concave function, its maximum on an interval is always attained at some point inside the interval.

Resulting explanation. As we have mentioned earlier, a person selects the effort e_0 for which the expression $r(e) - e$ attains its largest possible value.

Of course, people's abilities are not unbounded, there are certain limits within which we can apply the efforts. Thus, possible value of the effort e are located within some interval $[\underline{e}, \overline{e}]$. Thus:

- When we reward for good performance, the corresponding function $r(e)$ is convex, thus the difference $r(e) - e$ is convex, and therefore, the selected value e_0 coincides either with $\underline{e}$ or with $\overline{e}$. Thus, if we dismiss the case $e_0 = \underline{e}$ (when the reward is so small that it is not worth spending any effort), we conclude that $e_0 = \overline{e}$, i.e., the person selects the largest possible effort – which is exactly what we wanted to achieve.
- On the other hand, when we punish for mistakes, the corresponding function $r(e)$ is concave, thus the difference $r(e) - e$ is concave, and therefore, the selected value e_0 is always located inside the interval $[\underline{e}, \overline{e}]$: $e_0 < \overline{e}$. Thus, the person will not select the largest possible effort – which is exactly what we wanted to avoid.

This indeed explains why rewarding for good performance works better than punishment for mistakes.

Comments.

- What if we have both reward for good performance and punishment for mistakes, i.e., $r(e) = \text{const} + c_+ \cdot \max(0, e - e_0) + c_- \cdot \min(0, e - e_0)$? In this case, for $c_+ > c_-$, the function is still convex, i.e., we still get a very good performance, but if $c_- > c_+$, the function becomes concave, and the performance suffers. Thus, to get good results, reward must be larger than punishment.
- It is interesting to observe that the optimal rewarding function

  $$r(e) = r_+ + c_+ \cdot \max(0, e - e_0),$$

  in effect, coincides (modulo linear transformations of input and output) with the empirically efficient "rectified linear" activation function $r(e) = \max(0, e)$ used in deep learning; see, e.g., [1, 2, 4]. So, not only people learn better when we use this function – computers learn better too!

Acknowledgments

This work was supported in part by the US National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

References

[1] O. Fuentes, J. Parra, E. Anthony, and V. Kreinovich, "Why rectified linear neurons are efficient: a possible theoretical explanations", In: O. Kosheleva, S. Shary, G. Xiang, and R. Zapatrin (eds.), *Beyond Traditional Probabilistic Data Processing Techniques: Interval, Fuzzy, etc. Methods and Their Applications*, Springer, Cham, Switzerland, 2020, pp. 603–613.

[2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Leaning*, MIT Press, Cambridge, Massachusetts, 2016.

[3] D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus, and Giroux, 2011.

[4] V. Kreinovich and O. Kosheleva, "Deep learning (partly) demystified", *Proceedings of the 2020 4th International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence ISMSI'2020*, Thimpu, Bhutan, April 18–19, 2020.

[5] R. T. Rockafeller, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.