COMPLEXITY OF COLLECTIVE DECISION MAKING EXPLAINED BY NEURAL NETWORK UNIVERSAL APPROXIMATION THEOREM

Raul Trejo
Sistemas de Informacion
Division de Ingeneria y Ciencias
ITESM (Instituto Technologico de Monterrey)
Campus Estadoy de Mexico
Apdo. Postal 2, Modulo de Servicio Postal
Atizapan, Mexico 52926

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
email vladik@cs.utep.edu

Collective Decision Making is Important for Expert Systems

The main objective of an expert system is to automate expert's decision making.

Traditionally, expert systems started by formalizing the experience of a *single* expert. However, many important decisions are too difficult and/or too important for a single person to make; such decisions are made by a *team* (*group*) of experts. To simulate such decisions, it is, therefore, desirable to simulate collective decision making.

Collective Decision Making is Difficult to Describe

There are pretty accurate models for describing *individual* decision making, but, unfortunately, no models describe *collective* decision making as accurately.

Why is collective decision making so difficult to describe? At first glance, if several experts put their brains together, we should get the solutions that are better than the solutions that each of them can come up with on their own. Sometimes, the resulting solution is indeed much better. However, in many other situations, the collective solution is much worse than the individual one. This is a well-known fact: there are many examples of "designed by a committee" programming languages, political solutions, etc., that are clearly inferior to their individually designed alternatives. This strange behavior raises two fundamental questions:

- The first question is: Why is collection decision making imperfect?
- The second question is: How can we describe, in the general cases, this imperfect collective decision making?

The first question was solved many decades ago, when K. Arrow proved his famous impossibility theorem, according to which there is no consistent and reasonable general way to combine preference of different experts (see, e.g., [8]). In this paper, we will discuss the second question: how to describe the imperfect collective decision making.

For this description, let us first recall how individual decision making can be described.

Description of Individual Decision Making

To describe individual references (utilities), a special *utility theory* has been developed; see, e.g., [1, 8, 9, 10].

The mathematical formalism of utility theory comes from the observation that sometimes, when a person faces several alternatives $A_1, \ldots, A_n$, instead of choosing one of these alternatives, this person may choose a *probabilistic* combination of them, i.e., A_1 with probability p_1, A_2 with a probability p_2, etc. For example, if two alternatives are of equal value to a person, that person will probably choose the first one with probability 0.5 and the second one with the same probability 0.5. Such probabilistic combinations are called (somewhat misleadingly) *lotteries*. In view of this realistic possibility, it is desirable to consider the preference relation not only for the original alternatives, but also for arbitrary lotteries combining these alternatives. Each original alternative A_i can be viewed as a *degenerate* lottery, in which this alternative A_i appears with probability 1, and every other alternative $A_j \neq A_i$ appear with probability 0.

The main theorem of utility theory states that if we have an ordering relation $L \succ L'$ between such lotteries (with the meaning "L is preferable to L'"), and if this relation satisfies natural consistency conditions such as transitivity, etc., then there exists a function u from the set $\mathcal{L}$ of all possible lotteries into the set R of real numbers for which:

- $L \succ L'$ if and only if $u(L) > u(L')$, and
- for every lottery L, in which each alternative A_i appears with probability p_i, we have $u(L) = p_1 \cdot u(A_1) + \ldots + p_n \cdot u(A_n)$.

This function u is called a *utility function*. Each consistent preference relation can thus be described by its utility function.

From Individual Preferences to Collective Preferences

Let n denote the total number of original alternatives. Our goal is then to describe the collective preference between the probability combinations of these alternatives, i.e., between *lotteries*. Each lottery $L \in \mathcal{L}$ can be characterized by the vector $\vec{p} = (p_1, \ldots, p_n)$ formed by the probabilities of the corresponding alternatives.

Let E be the total number of participants in the decision-making group. For simplicity, we will describe each participant by a number $e = 1, \ldots, E$. Then, for e-th participant, the preference of a lottery $\vec{p}$ can be described by the value $u_e(\vec{p}) = \sum u_{ei} \cdot p_i$, where u_{ei} denoted the utility of i-th alternative for e-th participant.

One of the major applications of utility theory is to economic problems. For such problems, it is desirable to express the *monetary equivalent* of this utility, i.e., the amount of money $m_e(\vec{p})$ that would lead to the same utility value $u_e(\vec{p})$. This term may be somewhat misleading for non-economic situations (just like the word "lottery" is somewhat misleading), but it is a standard term in utility theory.

To estimate this amount of money, we must know how, for this participant e, the utility depends on the money, i.e., we must know this participant's dependence $u = M_e(m)$. Then, we can describe the monetary equivalent m_e of the utility $u_e(\vec{p}) = \sum u_{ei} \cdot p_i$ as the value for which $M_e(m_e(\vec{p})) = u_e(\vec{p})$, i.e., as a value $m_e(\vec{p}) = F_e(u_e(\vec{p})) = F_e(\sum u_{ei} \cdot p_i)$, where F_e denotes a function that is inverse to the function M_e.

Thus, we know, for e-th participant, the monetary value $m_e(\vec{p})$ of each lottery $\vec{p}$. If this participant was alone, this monetary value would describe his or her preferences: if for two lotteries $\vec{p}$ and $\vec{q}$, we have $m_e(\vec{p}) > m_e(\vec{q})$, then for this participant, the lottery $\vec{p}$ is better than the lottery $\vec{q}$.

How can we combine these individual preferences into a collective preference? For each lottery $\vec{p}$, each participant e gets a utility that is equivalent to this participant getting a monetary value $m_e(\vec{p})$. Thus, for the entire group, this lottery is equivalent to getting the sum of money $M(\vec{p}) = m_1(\vec{p}) + \ldots + m_E(\vec{p})$ and distributing it accordingly between E participants.

Let us show that this sum $M(\vec{p})$ provides us with a natural description of collective preference. Indeed, let $M(\vec{p}) > M(\vec{q})$ for two lotteries $\vec{p}$ and $\vec{q}$.

- Let us first consider the simplest case, when $m_e(\vec{p}) > m_e(\vec{q})$ for all $e = 1, \ldots, E$. In this case, for all participants, the lottery $\vec{p}$ is better than the lottery $\vec{q}$ and therefore, the group will also prefer $\vec{p}$ to $\vec{q}$.
- Let us now consider the general case of $M(\vec{p}) > M(\vec{q})$, when:
  - for some participants e, we have $m_e(\vec{p}) > m_e(\vec{q})$, but
  - for some other participants e, we have the opposite preference $m_e(\vec{p}) < m_e(\vec{q})$.

  Then, those participants who gain from choosing $\vec{p}$ (i.e., for whom $m_e(\vec{p}) > m_e(\vec{q})$) can donate some amount of money to compensate those who will lose from choosing $\vec{p}$. Namely, we can define the *balance* b_e of e-th participant as $b_e = m_e(\vec{p}) - [m_e(\vec{q}) + g]$, where we denoted $g = (M(\vec{p}) - M(\vec{q}))/E$.

  - A participant whose balance is positive donates the corresponding amount b_e to the compensation fund;
  - a participant whose balance is negative receives the amount of money $|b_e|$ from the compensation fund.

  One can easily check that the sum of all balances that correspond to all participants is equal to 0 and therefore, the total amount of donated money is exactly equal to the total amount of distributed money. As a result, we will have a new situation in which everyone's monetary equivalent $m'_e(\vec{p}) = m_e(\vec{p}) - b_e = m_e(\vec{q}) + g$ (utility plus or minus monetary compensation) is greater than the compensation $m_e(\vec{q})$ corresponding to choosing $\vec{q}$.

Thus, if $M(\vec{p}) > M(\vec{q})$, then the lottery $\vec{p}$ is preferablt to the lottery $\vec{q}$.

Collective Decision Making: General Description and a Question

In view of the above discussion, the lottery $\vec{p} = (p_1, \ldots, p_n)$ is preferable to the lottery $\vec{q} = (q_1, \ldots, q_n)$ if and only if $M(\vec{p}) > M(\vec{q})$, where

$$M(p_1, \ldots, p_n) = \sum_{e=1}^{E} F_e(u_{e1} \cdot p_1 + \ldots + u_{en} \cdot p_n). \quad (1)$$

This is a general description of the functions that describe collective decision making. A natural question is: which functions can be thus represented?

Taking Interval Uncertainty Into Consideration

In reality, all measurements are not 100% accurate. Therefore, a more realistic question would be: For a given accuracy $\varepsilon > 0$, which functions can be represented, within this accuracy, in a form (1).

In precise terms: Given $\varepsilon > 0$, describe all the functions $F(p_1, \ldots, p_n)$ that are ε-close to functions M of type (1), i.e., for which $F(p_1, \ldots, p_n) \in [M(p_1, \ldots, p_n) - \varepsilon, M(p_1, \ldots, p_n) + \varepsilon]$.

Enter Neural Networks

The expression (1) is exactly the expression of an input-output relation of a three-layer neural network, in which:

- the first (input) layer inputs the values $p_1, \ldots, p_n$;
- the second layer consists of E neurons, each of which computes the corresponding expression $F_e(u_{e1} \cdot p_1 + \ldots + u_{en} \cdot p_n)$; and
- finally, the third layer contains a single neuron that simply adds up all the outputs of neurons from the second layer.

It is known that such neural networks are *universal approximators*, in the sense that for every continuous function $F(p_1, \ldots, p_n)$, and for every $\varepsilon > 0$, there exists a neural network for which the corresponding function (1) is ε-close to $F(p_1, \ldots, p_n)$ [3, 4, 5, 6, 7].

The Resulting Answer to Our Question

Thus: *In collective decision making, it is possible to have, for an arbitrary function $M(p_1, \ldots, p_n)$, the ordering $\vec{p} \succ \vec{q}$ if and only if $M(\vec{p}) > M(\vec{q})$.*

From this answer, we can make two conclusions.

First Conclusion: Collective Decision Making Can Be Arbitrarily Complicated

Our first conclusion is somewhat pessimistic:

- We started with a general form (1) that looked like it describes *specific* functions that correspond to collective decision making.
- However, it turned out that an *arbitrary* continuous function can be represented in this form.

Thus, collective decision making can be arbitrarily complicated.

Second Conclusion: It Is Very Appropriate to Use Neural Networks to Describe Collective Decision Making

The second conclusion is optimistic: Since the general expression for collective decision making coincides with the expression for a three-layer neural network, it is natural use known methods of training neural networks (see, e.g., [2]) to analyze the collective decision making process.

ACKNOWLEDGMENTS

This work was partially supported by NASA research grant NCCW-0089, NSF grants EEC-9322370 and DUE-9750858, and by Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-95-1-0518.

The authors are very thanful to Věra Kůrková for valuable discussions. Her contribution is so big that we truly consider her our co-author. We also thank the anonymous referees for their helpful comments.

REFERENCES

[1] Fishburn, P.C.; *Utility Theory for Decision Making*, John Wiley & Sons Inc., New York, 1969.

[2] Haykin, S.; *Neural Networks: A Comprehensive Foundation,* Macmillan College Publishing Company, New York, NY, 1994.

[3] Hecht-Nielsen, R.; *Kolmogorov's Mapping Neural Network Existence Theorem*, Proceedings of First IEEE International Conference on Neural Networks, 1987, pp. 11–14, San Diego, CA.

[4] Hornik, K.; Stinchcombe, M.; and H. White, H.; *Multilayer Feedforward Neural Networks Are Universal Approximators*, Neural Networks, 1989, Vol. 2, pp. 359–366.

[5] Hornik, K.; *Approximation Capabilities of Multilayer Feedforward Neural Networks*, Neural Networks, 1991, Vol. 4, pp. 251–257.

[6] Kůrková, V.; *Kolmogorov's Theorem is Relevant*, Neural Computation, 1991, Vol. 3, pp. 617–622.

[7] Kůrková, V.; *Kolmogorov's Theorem and Multilayer Neural Networks*, Neural Networks, 1992, Vol. 5, pp. 501–506.

[8] Luce, D.R.; Raiffa, H.; *Games and Decisions, Introduction and critical survey*, John Wiley & Sons, Inc., New York, 1957.

[9] Myerson, R.B.; *Game theory. Analysis of conflict*, Harvard University Press, Cambridge, MA, 1991.

[10] Suppes, P.; Krantz, D.M.; Luce, R.D.; Tversky, A.; *Foundations of measurement. Vol. II. Geometrical, threshold, and probabilistic representations*, Academic Press, San Diego, CA, 1989.

References

[1] P. C. Fishburn, *Utility Theory for Decision Making*, John Wiley & Sons Inc., New York, 1969.

[2] S. Haykin, *Neural Networks: A Comprehensive Foundation,* Macmillan College Publishing Company, New York, NY, 1994.

[3] R. Hecht-Nielsen, "Kolmogorov's Mapping Neural Network Existence Theorem," *Proceedings of First IEEE International Conference on Neural Networks,* 1987, pp. 11–14, San Diego, CA.

[4] K. Hornik, M. Stinchcombe, and H. White, "Multilayer Feedforward Neural Networks Are Universal Approximators," *Neural Networks,* 1989, Vol. 2, pp. 359–366.

[5] K. Hornik, "Approximation Capabilities of Multilayer Feedforward Neural Networks," *Neural Networks,* 1991, Vol. 4, pp. 251–257.

[6] V. Kůrková, "Kolmogorov's Theorem is Relevant," *Neural Computation,* 1991, Vol. 3, pp. 617–622.

[7] V. Kůrková, "Kolmogorov's Theorem and Multilayer Neural Networks," *Neural Networks,* 1992, Vol. 5, pp. 501–506.

[8] D. R. Luce and H. Raiffa, *Games and Decisions, Introduction and critical survey*, John Wiley & Sons, Inc., New York, 1957.

[9] R. B. Myerson, *Game theory. Analysis of conflict*, Harvard University Press, Cambridge, MA, 1991.

[10] P. Suppes, D. M. Krantz, R. D. Luce, and A. Tversky, *Foundations of measurement. Vol. II. Geometrical, threshold, and probabilistic representations*, Academic Press, San Diego, CA, 1989.