# Nash Social Welfare, Logic, and Inductive Game Theory: An Application on the Russian invasion of Ukraine

by

Mamoru Kaneko

Emeritus Professor

University of Tsukuba, Waseda University

3rd April 2023

**Preface**: This lecture introduces the theory of *Nash social welfare* (NSW) *function,* and discusses its implications for peace and justice, particularly in the context of the Russian invasion of Ukraine. Though the NSW theory consists of extreme postulates from the viewpoint of extant social thoughts, it is becoming inevitable for the present world. As expected, it involves various conundrums at the foundational levels in the socio-economic and philosophical senses, in particular, the problems of small number of people versus a large number. I have been working on these problems for many years. Today, I discuss the NSW theory together with developments on related subjects, logic, and the inductive game theory, as well as their applications to the Russian invasion of Ukraine. Though I ignore many details, I would be happy to explain them at other occasions.

## 1. Beginning of My Research Career

My main research activity started with *Nash social welfare theory*, on which I worked with Kenjiro Nakamura. The results were published in Kaneko-Nakamura (1979a, '79b). The NSW function is given as

$$(1)\ \ w(u, x) = \sum_{i \in N} log[u_i(x) - u_i(x_0)],$$

where

- $N = \{1, \dots, n\}$ is the set of people in the entire world;
- $u = (u_1, \dots, u_n)$ is a list of (*Neumann-Morgenstern*) utility functions;
- $x_0$ is the worst state of the world, called the *origin*;
- $x$ is a candidate of alternative world states, to be evaluated using the NSW function.

The theory led me to various questions on its applications to practical/social problems as well as foundational questions on its ethical/normative status. The paper (Kaneko-Nakamura 1979a, p.423) assumed the *origin* $x_0$ as

> *the worst state for all individuals that we may imagine, e.g.,*
> *all the members of the world die.*

Around that time, I found Hobbes' (1651) *Leviathan*, and I was convinced that the NSW theory would be the right way to discuss world social justice. Around 1982, I discovered Einstein's principle of the *world peace*:

> *the objective of avoiding the total destruction of the world*
> *must have priority over any other objective.*

The total destruction of the world became feasible with the use of nuclear bombs during the 1940s−1950s, while Hobbes' state of nature was the practically worst case in the 17th century; the world of that time was infinite relative to the scope of people.

In the NSW theory, the total destruction corresponds to the worst state $x_0$. The origin $x_0$ is evaluated as $x = x_0$;

$$(2) \quad w(u, x) = \sum_{i \in N} log[u_i(x) - u_i(x_0)] = -\infty.$$

If a worldwide alternative $x$ includes the genocide of a group of people, the function $w(u, x)$ takes value $-\infty$, too. This $x$ is the same as the origin $x_0$ with respect to the NSW theory. I will examine the genocide in Ukraine committed by Russia, later.

At that time, my knowledge was limited, I was not able to have a systematic thought on the NSW theory. I did not understand Hobbes' social contractarian theory from his *state of nature* to the *absolutist state*. I raised the following naive questions.

**Question 1**: Because the NSW function is real-valued, its application to a social problem can be formulated as a simple maximization of $w(u, x)$. However, this is too simplistic. What kind of alternatives $x$ should the NSW function be applied to?

**Question 2**: Formally, the NSW theory is a different but equivalent formulation of Nash's (1950) bargaining solution (cf., Kaneko (1980)). The literal aim of the NSW theory is to study social welfare, rather than bargaining. Why do they have equivalent structures? What are the substantive differences?

In the 1980s, the problem of rationality became popular in the game theory community, due to Selten (1975). The studies in this area, however, examined the stability of an equilibrium and/or representations of beliefs/knowledge. It was developed in the direction of the Bayesian game theory, and did not directly address the problem of rationality. The term "rationality" is often treated as a synonym to "utility maximization". The concept of rationality must be the appropriateness of our thinking. We need to consider our thinking more directly. Gradually, I made up my mind to study logic.

Around 1984, I started studying logic seriously. I needed quite some time to start understanding what logic is, but finally I discovered that logic is intimately related to our intellectual activities in that symbolic manipulations express intended meanings.

I have a good example: One day, I observed a small child answered literally to the question "*How old are you?*" asked by some person:

> "*I am two years and three months old*"

Of course, he was unable to explain the meaning of his answer, but it was correct. If we ask ourselves whether we are different from that boy, we find that our knowledge is similar only with differences in degree.

The basic postulate of logic begins with the idea of symbolic operations, which is a radical treatment of bounded rationality. Now, Q1 can be answered; the detailed maximization of the NSW function could be meaningful only when the number of people $n$ is small, i.e., 2 or 3. When $n$ is large such as the population of a nation, the NSW function can be applied to a choice of an institution ignoring detailed differences among people. Other than institutions, an application of the NSW function offers only a choice for extreme alternatives such as whether to "stop genocide or not," neglecting almost all small details.

## 2. Bounded Rationality, Proof Theory, and Inductive Game Theory

Around 1985, I met a proof theorist, Takashi Nagashima, and started working on *game logic* with him. *Proof theory* deals with a pure form of logic. The papers with him (Kanko-Nagashima (1996,'97)) focused on a critique on logical thinking involved in game theory. Later, Nobuyuki Suzuki, Tai-Wei Hu, and I have developed epistemic logic with more emphasis on finiteness (cf., Kaneko-Suzuki (2003), Hu-Kaneko-Suzuki (2019)).

Around 1995, I started working on *inductive game theory* (IGT) with Akihiko Matsui and Jeffrey J. Kline (cf., Kaneko-Matsui (1999), Kaneko-Kline (2008)). This theory interprets symbolic pieces of information as a source for understanding in social contexts. This shares ideas of patterned behavior with "convention" of Neumann-Morgenstern (1944) and Lewis (1969).

The above developments raised the following question:

**Question 3**: We have appealed the calculation result (2) from the NSW function by referring to our intuition. What is this intuition? Or, does each of us have the source for the intuition such as morality in mind?

The same is asked for the Nash bargaining theory. His theory with two persons is understood as the result calculated by each logically rational player. The first question of Q3 is answered, in that, intuition is based on calculations by the logical ability of our mind based on the required rules such as the Nash axioms. This differs from the assumption that some morality is hidden in our mind.[1] Our logical ability is bounded, but when the situation is simple, such logical thinking works. Thus, the second question of Q3 was affirmatively answered.

Now, what is "morality" in our ordinary lives? This has two sides; one side is the answer to Q3. The other is social: morality has been emerging in human relations. IGT starts with this view; Kaneko-Matsui ('99) studied the relation between discrimination as behavioral patterns and prejudices as *ex post* rationalization. Once morality is accepted in society, people follow it as behavioral principle.

---

[1] In moral philosophy, it is implicitly presumed that morality is hidden in the minds of us, cf., Feldman (1978).

## 3.  Barbarity from Histories, and the Orwellian Denial of Truth

When I heard the news on the Russian invasion of Ukraine and the genocide there, I was shocked very much. Since then, Russian barbarity has persisted up to now. The question is:

**Question 4**: How do we understand Russian behavior from the viewpoint of the NSW theory, logic, and IGT?

Many assertions by Putin and his fellows have puzzled me; the words "Russia" and "Ukraine" are reversible in their assertions. These remind me of George Orwell's "*1984*" (1949), which describes the authoritative state and its severe surveillance together with language control. The following question can be raised:

**Question 5**: How do we understand the assertions of showing no respect for truth/reality?

Up to now, Putin has coherently shown that he wants to change the world state $x$ before 2022 into the new state $y$ so that the Ukrainian people are eliminated and Ukraine becomes a part of Russia. That is, the new utility levels are

$$(3) \quad u_i(y) \leq u_i(x_0) \text{ for Ukrainian person } i, \text{ but}$$

$$u_i(y) > u_i(x_0) \quad \text{for Russian person } i.$$

One basic postulate of the NSW theory is that all the people in the world want to avoid the origin $x_0$, which corresponds to Einstein's peace principle. This world alternative $y$ gives the same value $-\infty$ as the origin $x_0$. We ask the following questions.

**Question 6**: Do we count the people who show (3) as members of the world? Can such people exist in general? Is Putin exceptional?

The first question could be negatively answered by the constitution of the *World federal government* (WFG). It states that such people should be eliminated from the world, unless they change their behavior and thought to be compatible with the constitution.

The second is a purely *positive* scientific question, and it is rephased as: may such people/cultures be created in a world together with historical backgrounds? Russian history shows extremely barbaric incidents continuously under authoritarian regimes from its beginning. Such historical backgrounds have created moral values for the governing classes to treat other people as objects/animals.

The implication of IGT is that human beliefs and morality are products of historical social interactions, and beliefs are derived after observing social patterns of behaviors.

Finally, we return to Q5, which deals with the lack of respect for truth/reality. In a severe authoritarian regime, the value of survival is more important than truth. Even people's thoughts may become so controlled that it becomes impossible to think about truth or reality, which is well expressed as language control in Orwell's *1984*; for example, by eliminating the word "revolution," the nation would be free from revolution itself. People who have grown up under such circumstances give very little values to reality/truth.

## 4 Conclusions

I described the development of my research on the NSW theory. To understand the theory more deeply, I have worked on logic and IGT. I mentioned their applications and implications, as well as those of the NSW theory to understand the phenomena caused by the Russian invasion.

For long time, I got stuck on the idea of the *world government*, though I had vague ideas to deny this notion. Only a few years ago, I found that this should be replaced with the concept the *world federal government* (WFG). Kaneko (2019, revised 2021). This suggests that it consists of independent nations and should be governed by the constitution of the WFG, which should be coherent with the NSW theory, together with the restrictions mentioned in this lecture.

The NSW theory together with WFG forms a social contractarian theory for the entire world; its constitution should have various non-detailed statements such as "avoiding terrible incidents like genocide, famine, etc."

An implication of the Nash bargaining theory and the origin $x_0$ is that each person should have total independence and freedom, because each is given the integrity equally comparable to the entire world. This prohibits treating people as objects or animals.

Democracy and market economy are key institutions and tools for the practical management of the world. In Hobbes's time, neither democracy nor market economy was practiced. In the present world, they are indispensable, while having certain drawbacks. Majority rule is the foundation of democracy, but minority people may be discriminated against. Economics teaches that Pareto optimality is achieved through market economy, but it may lead to serious inequalities as well as worldwide problems such as global warming, cf., Hammond-Kaneko-Wooders (1989), Kaneko-Wooders (1994). We should study carefully constraints on democracy and market economy.

**Afterword:** I tried to convey my theoretical thought from 1976 to now. Specifically, I talked about integrity of each person and independence of each nation. These are directly applied to Ukraine, and they do not recommend an easy compromise. In the Ukrainian case, a cease-fire without going back to the state of 2014 gives a moratorium period to Russian, and after it, perhaps, Russia will start another invasion.

I heard that the Senate of SGH has started the procedure of awarding honorary doctorate to President Zelensky. I myself entirely support him and Ukraine from my sincere heart and theoretical thinking.

My theories help thinking about those basic problems of the present world, but they need more systematic thoughts from the foundations to practical applications. It is a long way to go. I finish this lecture with the quotation from Hippocrates 460~375 B.C.

*Life is short, the art long, opportunity fleering, experience unreliable, judgement difficult.*

## References

Einstein, A., (1968, original), Ideas and Opinions, New York, Bonanza Books.

Feldman, F., (1978), *Introductory Ethics*, Prentice-Hall, Englewood Cliffs.

Hammond, P., M. Kaneko, and M. H. Wooders, (1989), Continuum economies with finite coalitions: core, equilibria and widespread externalities, J*ournal of Economic Theory* 49, 113-134.

Hobbes, T., (1991, original 1651), *Leviathan*, Cambridge University Press. Cambridge.

Hu, T., M. Kaneko, and N.-Y. Suzuki (2019), Small infinitary epistemic logics, *Review of Symbolic Logic* 12, 702-735, doi:10.1017/S1755020319000029

Kaneko, M., (1980), An extension of the Nash bargaining problem and the Nash social welfare function, *Theory and Decision* 12, 135-148.

Kaneko, M., (2018), Exploring new socio-economic thought for a small and narrow earth, *Advances in Applied Sociology* 8, 397-421.

Kaneko, M., (2021, revised version of Kaneko (2018), Exploring new socioeconomic thoughts for a small and narrow world: unity and decentralization, New Horizons in Education and Social Studies Vol.11, Chap.1, 1-21. ed. Dr. Ritu Singh, Book Publisher International, London.

Kaneko, M., J. J. Kline (2008), Inductive game theory: A basic scenario, *Journal of Mathematical Economics* 44, 1332-1363.

Kaneko, M., and A. Matsui (1999), Inductive game theory: discrimination and prejudices, *Journal of Public Economic Theory* 1, 101-137.

Kaneko, M., and K. Nakamura (1979a), The Nash social welfare function, *Econometrica* 47, 423-435.

Kaneko, M., and K. Nakamura (1979b), Cardinalization of the Nash social welfare function, *Economic Studies Quarterly* 30, 236-242.

Kaneko, M., and T. Nagashima (1996), Game logic and its applications I, *Studia Logica* 57, 325-354.

Kaneko, M., and T. Nagashima (1997), Game logic and its applications II, *Studia Logica* 58, 273-303.

Kaneko, M., and N.-Y. Suzuki, (2003), Epistemic models of shallow depths and decision making in games: Horticulture, *Journal of Symbolic Logic* 68, 163-186.

Kaneko, M., and M. H. Wooders (1994), Widespread externalities and perfectly competitive markets: Examples, *Imperfection and Behavior in Economic Organizations,* eds. R. Gilles and P. Ruyes, Kluwer Academic Publisher, 71-87.

Lewis, D. K., (1969), *Convention; Philosophical Study*, Harvard University press, Cambridge.

Nash, J. F., (1950), The Bargaining problem, *Econometrica* 18, 155-162.

Orwell, G., (1949), *Nineteen Eighty-Four*, new American library. New York.

Selten, R., (1975), A reexamination of the perfectness concept for equilibrium concept in extensive games, *International Journal of Game Theory* 4, 25-55.

Von Neumann, J., and O. Morgenstern, (1944), *Theory of Games and Economic Behavior,* Princeton University Press, Princeton.