Modelling of Non-commuting Measurements

V. I. Danilov (Central Economics and Mathematics Institute, Russian Academy of Sciences, Moscow)

Abstract

In the paper we propose an approach which allows to discuss on formal level non-commuting measurements. Such measurements 'disturb' a measured system and change its states. The key structure used in the description of such measurements is the structure of ortho-poset (which is a Boolean lattice in the classical case of commuting measurements). States of the system are realized as probabilistic measures on the ortho-poset. We give an application of the proposed approach to decision-making under 'non-classical' indeterminacy and to modelling 'non-classical' preferences.

Key words: first-kind measurement, inconsistent measurements, event, ortho-poset, transition probability space.

JEL classification: C44, D81

1 Introduction

Measurements (widely understood as observations, tests, experiments, polls, etc.) underlie the study of practically any system. **Measurements provide** primary information on the system under study, its properties, possible interrelations between the properties and the dynamics of these properties. The other important part of science is comprehension of the observations, their systematization, construction of a theory. The conclusions of the theory are usually verified by experiments again. Certainly, the second part of science seems more interesting and creative from the intellectual point of view. And yet it is only a superstructure that rests on the basis of measurements and often ambiguously determined by this basis.

Measurements are performed with measuring instruments which display the measurement results. Here is for example a thermometer: we submerge it in the medium to be measured and note the height of the mercury column against its scale; or we poll a respondent. Measurements are often performed for some practical rather than scientific purposes. You want to know the temperature outdoors to decide what to put on. Or you play roulette and get one or other payoff depending on the outcome.

Certainly, a real measurement is always taken with errors caused by imperfection of the measuring instrument and by its effect on the measured object. However, measuring instruments became both more precise and finer over time. This progress in measurement technology has led to the notion of an ideal measurement, ideal both in the sense of precision and in the sense of being intangible for the object of measurement ["with fingers light as sleep..."]. A by-product of this idealization is the notion of properties. A stable result of a measurement points to the presence of a property (color, mass, temperature). As often occurs, the relationship between the measurements and the properties has come to be turned inside out. The popular belief is that properties of things exist irrespective of and before any measurements; measurements only reveal them for us.

But if measurements do not change properties, we can perform various measurements in any order and utterly obtain comprehensive (exhaustive) knowledge about all properties, that is, about the state of the system under study. It only remains to understand the dynamics (evolution law) of the change in the states of the system, either single or interacting with others. We call this gnosiologically optimistic picture the classical one by analogy with classical mechanics in which it was most distinctly embodied.

Quantum mechanics struck a blow at this picture. The Heisenberg uncertainty **principle** asserts that there are quantities which cannot be simultaneously measured to whatever accuracy. However, it is better to speak about non-commuting or inconsistent measurements, that is, about measurements whose results depend on the measurement procedure. It is not accidental that the problem of measurements occupies a central place in quantum mechanics (see Quantum Theory and Measurement (1983)). In principle, however, inconsistent measurements can refer not only to phenomena of the micro-world.

Quite a lot of data resembling 'paradoxical' phenomena of the micro-world (like interference) and hinting at inconsistent measurements have recently appeared in the humanities; hinting, because measurement errors are still too large here. **We** give only one example, the *disjunction effect* from the work Tversky, Shafir (1992).

Let us imagine that a student is asked before the examination whether he/she is going to Hawaii for a holiday. The answers are mainly negative. After the examination (whether passed or failed) the wish to go to Hawaii is appreciably stronger. The intermediate measurement, the examination, has changed the result of the first measurement, the poll.

It is not difficult to find an explanation for this paradox at the common-life level. If the student has passed the exam, he is in good mood, wants to reward himself and to celebrate the success. If he has failed, he wants to relax and forget the failure. The reasons for the trip are different, and before the examination his emotional mood did not point to any of them. Therefore, there is no impetus for him to to go to Hawaii. Other paradox origination mechanisms are possible as well. But there is an obvious parallel with the well-known quantum-mechanical experiment with two slits. When one slit in the screen is open the probability that a particle (electron, photon, ...) hits the detector is fairly high; when both slits are open, the probability for a particle to hit the detector drastically decreases.

In short, non-commuting measurements can occur not only for microscopic systems but also for 'large' systems. There appear more and more works dealing with these 'quantum-like' phenomena in economics, sociology, psychology, linguistics, etc.; some idea of this can be gained from the proceedings of the 'Quantum interactions' conferences (2008). This raises a question of elaborating a theory of non-commuting measurements. An example of such a theory (**the** so-called *Hilbert space model*) for the needs of quantum mechanics was proposed by von Neumann (1932). It is not surprising that on facing 'quantum-like' phenomena many researchers (Franco (2008), Lambert-Mogiliansky, Zamir, Zwirn (2004)) tried to apply directly the quantum mechanics formalism without taking the trouble to explain why it is in place here. An example of this 'search under the lamp' is the work by Franco (2008). The author proceeds from the fact that experimental science (in social field) has accumulated a lot of paradoxical facts in recent years. Therefore, "quantum mechanics, for its counterintuitive predictions, seems to provide a good formalism to describe puzzling effects of contextuality". Yet, Wright (1978) warned that "Hilbert spaces, and Physics in general, are too tightly knit to serve as a criterion for less highly structured situations such as those occurring in social sciences. Experiments in these fields have long been demonstrating these phenomena as the effect of one measurement upon another, and we therefore should expect nonclassical states to be a rule rather than an exception".

Later Wright returned to this thought in his work Wright (2004) devoted to comparison of quantum mechanics and social sciences and stressed that "social science is even more nonclassical than quantum mechanics"; and that it would be good to have a theory more general than the Hilbert space model. An outline of this general theory based on the joint works with A. Lambert-Mogiliansky (Danilov, Lambert-Mogiliansky (2008a, 2008b)) is proposed below. It is worth noting three characteristic features of the quantum-like behavior which are intimately related to one another:

1. Unavoidably probabilistic character of predictions.
2. Existence of inconsistent measurements.
3. Influence of measurements on the state.

The first characteristic feature does not require any particular explanations, being commonly known. What one primarily points out in speaking about quantum mechanics is the loss of determinism. Of course, the probability exists in classical mechanics too (as in other sciences of nature and man), it is no news. However, in the classical science the probabilistic character of answers was attributed to an incomplete knowledge of the state of the system. When a cannon is fired, the shell hits a random place. But this is because we do not exactly know the direction of the barrel, weight and composition of gunpowder, wind force, etc. The situation is similar for coin tossing or stock pricing. However, if we exactly knew the state of the system, it would be in general possible to judge accurately
✓ the result of any measurement. Let us assume that we deal with two **measurements** A and B. Initially, we may be unaware of the results. After performing measurement A we obtained a result a. Having performed measurement B, we got b. Now we exactly know the results of any future measurement A and B.

Previous reasoning implicitly suggested that the performing of measurement B did not change the result of the previous measurement A, and vice versa. In this case **the**
✓ **measurements** A **and** B **are called** *consistent* (or ***commuting***); it does not matter in what order the measurements are performed; roughly speaking, they can be performed simultaneously. The situation is quite different if our measurements A and B are inconsistent. In this case it matters whether we first perform measurement A and then B or vice versa. With inconsistent measurements, the previous reasoning about elimination of uncertainty does not hold any longer. Performing measurement A and obtaining result a, we eliminate the uncertainty about measurement A. However, we do not know what measurement B will yield. Having performed B, we eliminate the uncertainty about B.

Yet, now **we** have no reasons to believe that measurement A will yield the previous result a. We are again uncertain about A.

But why on earth the repeated measurement A may yield a result different from a? A somewhat tautological answer is that measurement B changed the previous state of the system. The previous state yielded the result a for measurement A; the new state yields different (now random) results. Consequently, the state has changed. Thus we come back to the first line—unavoidable uncertainty. Eliminating the uncertainty about one measurement, we increase the uncertainty about the other measurement. And this goes on endlessly. While in the classical approach (or for the classical system) measurements purify the state more and more, ultimately making it free of dispersion and completely known, the nonclassical system attains the pure state, and further measurements only turn some pure states into others. Having reached the level of maximum (though incomplete) information, we are then simply sliding along this level.

In a word, observation of a system may lead to a change in the state. In particular, it may happen that the measured property simply **did** not exist before the measurement and may in a sense be produced by the measurement itself. That is what Kahneman and Tversky, who carried out lots of experiments to reveal preferences and limited rationality, write (Kahneman, Tversky (2000)): “The classical theory of preferences assumes that each individual has a well-defined preference order... But the observed preferences are not simply read off from some master list; they are actually constructed in the elicitation process”[1]. In other words, a person could have no preference at all so that it is the question itself which makes him fall to thinking and solve the problem for himself. Again Wright (2004): a person is asked whether he is angry. The answer “Yes” may move the person from a **blended** emotional state to an **experience of** angry and change his answers to any other questions.

The system of notions proposed below is not based on nothing, of course. The first attempt to theoretically **justify** the Hilbert space model was made in the work Bikhoff, von Neumann (1936). Together with the later work Mackey (1963) this gave an impetus to development of the so-called Quantum Logic. Its status can be learned from Beltrametti, Cassinelli (1981), Holland (1995), and Coeke, Moore, Wilce (2001). The major difference of our work from ‘quantum logic’ is that we more explicitly pose the question of the effect of measurement on the state, the question of ‘updating’.

Here is a brief review of what follows. We begin with three examples of inconsistent measurements. The notion of measurement is more formally discussed in Section 3. Then we dwell upon the notion of ‘properties’ or ‘events’. This allows the first application of the developed formalism to the decision making problem. In Section 6, the central notion of state is discussed and the definition of the model of the measured system is introduced. The model of nonclassical preferences is presented for illustration. Finally, we consider the model of transition probabilities which can be thought of as non-Abelian generalization of the Hilbert space model. In the Conclusions, possible causes for inconsistency of measurements are discussed.

[1] In physics this view comes more difficult if Peres still has to explain that “quantum theory is incompatible with the proposition that measurement discovers some unknown but pre-existing property”.

2 Examples of non-commuting measurements

Before we begin the discussion of such fundamental notions as **measurements and states**, it is worth **to consider** three simple stylized examples of inconsistent measurements; stylized in the sense that nobody **has** performed measurements exactly like this. All three examples are formally equivalent to one another though they refer to different fields and are convincing to a different extent.

1. The first example is taken from quantum physics. Atoms and elementary particles have a magnetic moment. Stern and Gerlach invented an apparatus with which they measured the magnetic moment of silver atoms (for details, see Feynman, Leyton, Sands (1965)). A beam of particles passed through an inhomogeneous magnetic field generated by a magnet with a sharpened pole; the presence of the magnetic moment in the particle caused deviation of the particle along the gradient of the change in the magnetic field. Let us imagine that the particles move away from us and that the magnetic field gradient is directed vertically. Then the particles will deviate up or down; the amount and direction of the deviation depend on the magnitude of the magnetic moment of the particle and the angle between the magnetic moment of the particle and the vertical.

Since magnetic moments of particles are originally randomly oriented, deviations will be more or less uniformly distributed over some vertical segment. What a surprise it was to find that particles always deviated by the same value (up or down). That is, magnetic moments of these particles were always oriented either strictly up or strictly down. But how could the particles know that they would be measured by an apparatus oriented vertically and not in any other direction?

If one now takes only the half of the initial beam which deviated upward and pass it through another Stern–Gerlach apparatus with the same orientation, this beam will not be split. All its particles will deviate upward. As Feynman says, the first apparatus produced a beam of 'purified' objects: all atoms were in a particular state within it. Their behavior in the apparatus of the same type is determined and predictable.

However, let us pass this filtrated beam through the Stern–Gerlach apparatus now oriented horizontally. We will find out that the beam (even purified vertically) is split again into two beams deviating to the left and to the right. If we take any of these subbeams and pass them through the Stern–Gerlach apparatus oriented vertically, we will get up–down splitting again, though we started with a pure beam of particles whose magnetic moments were directed upward.

It is a **classical** example of inconsistent measurements. Why do particles behave in such a strange way? Physics refuses to answer this question.

2. The second (now quite speculative) example refers to **a** man. An individual under test is asked to answer the question V which allows only two answers, U and D. Upon thinking, he gives the answer U. If he is asked again, the answer will be the same. This appears to indicate a property of his; let us call it U. Well, now he is asked another question H with the possible answers L and R. He answers R, and answers this way repeatedly. But we return to the previous question V and ask it again. And he answers D!

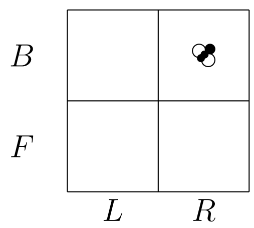
The situation is similar to the example with Hawaii in the Introduction. The first question is whether he plans to go to Hawaii. The second measurement is the examination. And they are inconsistent.

What this situation can arise from? Above we have already given an 'emotional' explanation. There can be another one related to the limited capacity of the human memory. Imagine that our individual has only one memory cell. On answering (for himself) the first question V, he writes the question on a piece of paper and put it into the cell. If he is asked the same question again, he takes the written answer from the cell and simply reads it. But if he is asked the question H, the previously written answer will be useless. It will be necessary to look for the answer to this question H, which is again written on a piece of paper and placed in the memory cell. If we now return to the first question V, the written answer will be useless again, and again it will be necessary to think the question over, and the new answer may be different from the initial one.

Of course, this explanation (like the example itself) looks rather far-fetched. The human memory does not consist of a single memory cell, and constancy in answering is doubtful. And yet it at least partially hints at possible inconsistency of some measurements of 'large' systems.

3. In the first example we had a distinct effect but no clear explanations. In the second example we had a weak effect and vague explanations. The more interesting is the third example, where everything is ultimately clear.

A box with a fly. Imagine a square box divided into four rooms by two opaque partitions. ? In the box there is a **permanently** flying fly. We are interested in the location of the fly. The figure below schematically shows the top view of the box in question.



It is assumed that there are only two ways to find out something about the location of the fly. One (let us call it measurement H) is to find whether the fly is in the left-hand or right-hand half of the box. To this end, a partition (horizontal) is lifted and thus the rooms RF and RB are united and can be viewed all the way through. The same occurs with the rooms LF and LB. As a result, we get the opportunity to see (viewing from the front or from the back, it does not matter) what is in these 'double' rooms and, in particular, to understand in which half, right-hand or left-hand, the fly is. After the observation the partition is lowered down and the fly cannot be seen again. The second measurement V makes it possible to find out whether the fly is in the front or back half of the box. To this end, the (vertical) partition is lifted and, viewing from the side, we see in which half, forward or backward, the fly is. After that the partition is lowered down again.

Thus, we have two measurements, the first one H with the results L and R, and the second one V with the results F and B. Initially, it is rather obscure to us what occurs if we perform the first measurement; the results L and R seem if not equiprobable then quite possible. On carrying out this measurement we obtained the result, say, R. That is, the fly is on the right. Note that if we repeat this measurement H, we will get the same result R. And even if we repeat this measurement three, five, or ten times, the result will be R. So this is not a fiction but a reliable knowledge. Well, now let us perform the second measurement V. Assume that it yielded the result F: the fly is in the front. However, we lifted the 'vertical' partition, and the fly could fly from the room FR to the room FL (and back again, and even several times). Thus, when we repeat the first measurement H, we may well obtain the result L.

This simple example demonstrates three above-mentioned features of quantum-like behavior (though it is not a mystery, and everything is utterly clear). We always has an uncertainty; knowing exactly the result of one (last) measurement, we can only probabilistically predict the result of the other measurement. The measurement explicitly changes the state ('actual location of the fly') by virtue of the fact that a partition is lifted during the measurement and the fly can fly from the current room to the adjacent one. Finally, our two measurements H and V are inconsistent. They cannot be performed simultaneously; one of them, when performed, changes (makes uncertain) the result of the other, as if **cancels** it.

Now let us move on to a more detailed discussion of measurements.

3 Measurements

What is a measurement? We bring the system under study into (temporal) interaction with **a** measuring instrument. This interaction results in some changes occurring in the instrument (and in the measured system). Some of these changes are directly visible and are thought to be the result of the measurement. For example, measurement of temperature by a thermometer: the mercury column rises and stops at a mark on the scale. Everybody can give lots of other examples. In a word, the most important component of the measuring instrument is the *scale* on which we read the ***outcome*** of the measurement. The measuring instrument is denoted by M, and $O(M)$ indicates a set of possible **outcomes** of the instrument. These are most often numbers (**of** grams, meters, seconds, degrees), but $O(M)$ can in principle be a rather arbitrary set. For example, it can be the answers YES and NO to a particular question. For simplicity, in what follows we consider $O(M)$ to be a small finite set whose elements are unambiguously different from one another.

Thus, measurements are not arbitrary interactions but rather interactions yielding **an outcome, an** information, if you like. But this is not the only requirement on the measuring instrument. Another important property is repeatability, reproducibility of the result obtained. Assume that we have performed a measurement and have obtained **an outcome** a. Let us immediately repeat the measurement. (Needless to say that this implicitly suggests a nondestructive measurement. In physics, the measurement of the location of a photon results in disappearance of the photon. Or, they say, the following witch revealing method was practised in the Middle Ages. A suspected woman was thrown into water, and if she survived, it was interpreted as the Devil's protection.) If the same **outcome** a as in the first measurement is always obtained, one speaks (thanks to Pauli's good graces) about measurements of the *first kind*.

What could interfere this “first-kindness”? Three causes come to my mind. First, the instrument works with noise and yields though different yet similar results. We will leave this cause out considering that the instrument is ideal in terms of precision. Second is the dynamics of the measured system. If we measure **a location** of a flying plane, the repeated measurement will obviously differ from the previous one. Here everything is also clear; generally, we should make a correction for dynamics. The third cause is randomness of the measured value. Imagine that our system is a population; the measurement consists in taking a random representative of the population and measuring his height. It is clear that in the second measurement we will pick up a different representative and the measurement result will be different. In social sciences measurements like this are rather a rule.

In short, **the** first-kindness is an important and by no means obvious property of the measuring instrument. In other words, to make a good measuring instrument is a quite nontrivial and creative task (remember Stern–Gerlach apparatus). Measurements in classical physics were tacitly assumed to be of the first kind, and this property was not discussed at all; it appeared only in quantum mechanics in connection with importance of the process of observation and measurement. In social sciences, as far as I understand, the property of being of the first kind is not so much comprehended. For example, when a person is asked what he likes best, apples or bananas, and he answers, he will hardly be asked this question again. But why? It is perhaps subconsciously assumed that the tested person will repeat his answer. It is most likely to be that way, but it generally requires verification as well. In this connection, I like very much one **sentence** in the paper Busemeyer, Wang (2007). Speaking about the probabilistic nature of the choice and the random utility models, they point out, “... Yet there is something unusual about the probabilistic nature of choice by humans that is not captured by these theories. Good experimenters know that if you present the same choice problem back to back, without any filler problems, then choice behavior is surprisingly deterministic: people simply choose the same as before... Choice becomes probabilistic only when a problem is repeated with fillers in between repetitions.” Actually, it is said here that the act of choice is a first-kind measurement and that there exist inconsistent choice problems. It is a pity that these are but words and not clear-cut experimental facts.

Next, we will only deal with the first-kind measurements. Two (first-kind) measurements A and B (more precisely, two measuring instruments) are called *consistent*, or commuting, if in any series **consisting only of** these measurements the **outcomes** will be made up of the corresponding pair of **outcomes**. For example, if we carry out a series $AABABA$, the results should have the form $aababa$, where $a \in O(A)$ and $b \in O(B)$. In other words, performance of intermediate measurements by the instrument B between the measurements by the instrument A does not change the results of the measurement A. A similar definition can be given to consistency of three and more measurements.

If we have two consistent measurements A and B, we can produce a new, compound measurement. For example, like that: first we perform A, then B, and write down the results as a pair (a, b). We obtain a new, finer measurement, which is the first-kind measurement again. Repeating this operation, we arrive in the limit at the finest measurement (these measurements are called *full*). In principle, there can be two possible cases, where the unique full measurement exists and where several full measurements (which are already inconsistent with one another) exist. In the former case we get a classical picture. In the latter case the classical picture is impossible, and what picture is needed will be discussed a bit later. And now we will speak about the notion of properties (if one holds a 'passive' research position) or events (if the things are viewed ✓ from **an** 'active' decision-making position).

4 Events

Let us play with measurements following the work Danilov, Lambert-Mogiliansky (2008b). ✓ **A *pre-event* is** a pair (M, A), where M is some measurement and A is a subset in $O(M)$. In terms of content it means that we performed the measurement M and found that the outcome of this measurement belonged to the set A of possible outcomes. Why 'pre-'? Because the same event or property can be detected by different measuring instruments. For example, the pre-events $(M, O(M))$ and $(M', O(M'))$ are the same trivially **certain** ✓ event. And we wanted the referencing to measuring instruments to be avoided, wherever possible. How can it be done? Let us introduce the following

Definition. The pre-event (M, A) *entails* the pre-event (M', A') (we will designate ✓ this as $A \Longrightarrow A'$) if the measurement M' performed immediately after (M, A) **certain** yields the outcome belonging to A'.

We will assume fulfilled the following two axioms:

1. $A \Longrightarrow A$.
2. If $A \Longrightarrow A'$ and $A' \Longrightarrow A''$ then $A \Longrightarrow A''$.

✓ Note that axiom 1 is simply **the** first-kindness. Axiom 2 is less obvious. If we insert a measurement M' between the measurements M and M'' and do not pay even the slightest attention to its result, we obtain A''; why is A'' also obtained when we do not perform the intermediate measurement M'? Axiom M2 is mainly justified by the fact that it is fulfilled in the classical case (in some more general cases as well, see Section 8).

We say that two pre-events (M, A) and (M', A') are *equivalent* if $A \Longrightarrow A'$ and $A' \Longrightarrow A$. Classes of equivalent pre-events are called *events*. Axiom 2 results in that the relation $\Longrightarrow$ induces a partial order relation on the set $\mathcal{E}$ of events; it is designated as $\leq$. The ✓ class of **certain** pre-events $(M, O(M))$ is designated as $\mathbf{1}$; $\mathbf{0}$ denotes the class of impossible events $(M, \emptyset)$.

The poset (partially ordered set) $\mathcal{E}$ has another important structure – the *orthocomplementation* structure. Let (M, A) be a pre-event. The *opposite pre-event* will be represented by $(M, O(M) - A)$. Let us introduce one more axiom

M3. If $A \Rightarrow B$, then $\overline{B} \Rightarrow \overline{A}$ (where the over-bar denotes the opposite pre-event).

It follows from this axiom that pre-events opposite to equivalent pre-events are also equivalent, which allows speaking about opposite events. For the event E the opposite event is designated as $E^{\perp}$. The operation $\perp : \mathcal{E} \to \mathcal{E}$ is antimonotone, involutory, and !!! This is my mistake $E \vee E^{\perp} = \mathbf{1}$ for any event E. (In what follows $\vee$ denotes the **least** upper bound; this operation does not always exist. In this case the equality $E \vee E^{\perp} = \mathbf{1}$ simply means that the only event which $\geq E$ and $E^{\perp}$ is $\mathbf{1}$.) Posets with **such an** operation $\perp$ are called *orthoposets*. The elements a and b in the orthoposet are considered *orthogonal* (we write $a \perp b$) if $a \leq b^{\perp}$ (or, which is the same, $b \leq a^{\perp}$).

Let us introduce one more notion. The family $(a(i), i \in I)$ of the elements of the orthoposet $\mathcal{E}$ is called the *orthogonal decomposition of unity* (ODU) if the following two conditions are fulfilled:

1) For any $J \subset I$ there exists a supremum $\vee_{j \in J} a(j)$ designated as $a(J)$;

2) $a(I - J) = a(J)^{\perp}$.

In particular, all events $a(i)$ are mutually orthogonal and $a(I) = \mathbf{1}$. This justifies the term ODU.

Let M be a measurement, and let the family of subsets $(A(i), i \in I)$ constitute **a** partition of the set of outcomes $O(M)$ of this measurement (that is, $(A(i)$ do not intersect and upon joining give all $O(M)$). It is easy to understand that the corresponding events $a(i)$ (represented by the pre-events $(M, A(i))$) **form an** ODU. **An** ODU is called *admissible* if it is produced by some measurement. It is easily understood that for any event E the pair $(E, E^{\perp})$ is an admissible ODU. Though this is not much justified in the general case, it is convenient to believe that any ODU is admissible. In this case we can (with a reservation to be discussed in Section 6) forget about measurements and only deal with the orthoposet $\mathcal{E}$. However, this assumption is not so much innocent. It implies that we have an opportunity to construct measuring instruments in the cases where this is not obviously hampered.

Having an orthoposet of event, **one hardly resists** the temptation to tell how classical Savage's theory of decision making under uncertainty is carried over to the case of inconsistent measurements.

5 Theory of expected utility

Recall how the theory of decision making (more exactly, of risk bet estimation) applied to the classical case is constructed in Savage (1954). To avoid digressing to issues unrelated to the subject of interest, we slightly simplify Savages's axiomatics. Savage begins with the set S of states of 'nature' and treats an event as an arbitrary subset in S. A *bet* (or *action* in Savage's terminology) is an arbitrary function $f : S \to \mathbb{R}$. The intuitive meaning is as follows: if the state $s \in S$ has been realized, the player gets $f(s)$ rubles (more exactly, utils). The set of all bets is the space $\mathbb{R}^S$ (for simplicity, we consider the set S to be finite). Savage suggested that a weak-order relation ('preference') was given on $\mathbb{R}^S$. We will hold that this preference is given by the functional $CE : \mathbb{R}^S \to \mathbb{R}$; the bet f is not worse than the bet g if $CE(f) \geq CE(g)$. We can normalize CE by the condition $CE(1) = 1$ without loss of generality. Now let us impose two natural conditions on CE:

Monotonicity: if $f \geq g$, then $CE(f) \geq CE(g)$, and

Additivity: $CE(f + g) = CE(f) + CE(g)$.

It is easy to deduce from them that the functional CE is actually linear, and therefore it is given by the formula $CE(f) = \sum_{s \in S} f(s)\mu(s)$, where $\mu(s)$ are some numbers. From the condition of monotonicity the numbers $\mu(s)$ are nonnegative, and from normalization

$\sum_s \mu(s) = 1$. Therefore, the set of numbers $(\mu(s), s \in S)$ can be treated as a probability on S and $CE(f)$ as the expected value (utility) of the bet f. In other words, in order to estimate 'correctly' arbitrary bets, a person should ascribe subjective[2] probabilities to states of nature (and to arbitrary events as well, in view of additivity).

Let us have a look at how this scheme changes when we abandon the classical approach. As we will see, the changes are minim**al**.

1. First of all, we replace the Boolean algebra of events 2^S with the orthoposet $\mathcal{E}$ of events, as in the previous section.

2. The notion of **a** bet undergoes the most important change. Now **a** bet is made on the result of the measurement. **A** *bet* is a pair (M, f), where M is an admissible measurement and $f : O(M) \to \mathbb{R}$ is **a** function specifying the payoff in relation to the result of this measurement. Since we have agreed above to identify measurements and ODUs in $\mathcal{E}$, **a** bet is a pair (α, f), where $\alpha = (a(i), i \in I(\alpha))$ is **an** ODU and $f : I(\alpha) \to \mathbb{R}$ is **a** payoff function. The ODU α is called the *basis* of the corresponding bet. Bets with the basis α constitute a vector space $\mathbb{R}^{I(\alpha)}$, and the set $Bet(\mathcal{E})$ of all bets is **the disjoint union** join $\coprod_\alpha \mathbb{R}^{I(\alpha)}$ of these spaces; the **union** is taken over all ODUs.

3. The preferences of the decision maker (DM) are given on $Bet(\mathcal{E})$; as in the classical case, they are given by the functional $CE : Bet(\mathcal{E}) \to \mathbb{R}$. As previously, this functional is assumed to be normalized in the following sense. Let 1_α denote the bet (on the basis of the ODU α) which gives one util irrespective of the outcome. Then $CE(1_\alpha) = 1$.

4. As previously, we assume that CE satisfies two axioms: additivity and a version of monotonicity. Additivity refers to bets with the same basis and is formulated as before. As to the other axiom, in the general case it relates together bets made on the basis of different ODUs or measurements. Roughly speaking, it states that if one bet (α, f) always gives a larger payoff than another bet (β, g) (i.e., dominates it), its CE is higher.

But what does it mean that one bet dominates **another**? We say that **a** bet (α, f) *dominates* **a** bet (β, g) if $f(i) < g(j)$ **implies** that the events $a(i)$ and $b(j)$ are orthogonal. In other words, if the event $a(i)$ does not rule out the event $b(j)$, then $f(i) \geq g(j)$. In terms of content it can be interpreted as follows: whatever the result of the measurement α, if we immediately after that perform the measurement β, the payoff of the second measurement does not exceed the payoff of the first one. Conversely, if we perform the measurement β and after that we perform α, the payoff of the measurement α will not be lower. These properties allow us to hold that domination entails preference and to consider the following formulation of the monotonicity axiom as reasonable:

Monotonicity. If (α, f) dominates (β, g), then $CE(\alpha, f) \geq CE(\beta, g)$.

5. Now we approach the main issue: how can the functionals CE satisfying the additivity and monotonicity axioms (we will call them *normal* for brevity) be **constructed**? Do they exist, and, if so, how many of them are there? To answer these questions, we begin with **an** explicit **construction** of normal CE, the **construction** that is called *expected utility.* But in order to speak about expectations, one should have the notion of probability. Let us put in brief how this notion is formulated for an arbitrary orthoposet.

Definition. **A** *probability* (or *probabilistic measure*) on the orthoposet $\mathcal{E}$ is **a** mapping

[2]according to Savage, de Finetti, and others of the kind, there are no other probabilities at all.

$\mu : \mathcal{E} \to \mathbb{R}$ which is monotonic ($a \leq b$ entails $\mu(a) \leq \mu(b)$) and which possesses the following property: $\sum_i \mu(a(i)) = 1$ for any ODU $(a(i), i \in I)$.

It is easy to understand that then $\mu(\mathbf{0}) = 0$, $\mu(\mathbf{1}) = 1$, and for any $a \in \mathcal{E}$ $0 \leq \mu(a) \leq 1$ and $\mu(a) + \mu(a^{\perp}) = 1$ hold. Let us denote the set of all probabilistic measures on the orthoposet $\mathcal{E}$ by $\Delta(\mathcal{E})$. Since the convex mixture of probabilistic measures is the probabilistic measure again, the set $\Delta(\mathcal{E})$ is convex.

If the orthoposet $\mathcal{E}$ is a Boolean lattice, we get the classical notion of probability. If the orthoposet $\mathcal{E}$ meets an additional condition (orthomodularity; we will not formulate exactly this notion here and refer the reader to Beltrametti, Cassinelli (1981), Coeke, Moore, Wilce (2001), Holland (1995)), probability is additive in the following sense: if the event $a \vee b$ is a supremum of two ORTHOGONAL events a and b, then $\mu(a \vee b) = \mu(a) + \mu(b)$. In the case of orthomodular posets, probability can be defined as a nonnegative and additive function on $\mathcal{E}$ normalized by the condition $\mu(\mathbf{1}) = 1$. Thus, for orthomodular posets, **the** probability theory is maximal close to the classical (Boolean) case. Note, however, that there can be too 'few' probabilistic measures. In Danilov, Lambert-Mogiliansky (2008b) we give an example of the orthomodular poset with a sole probabilistic measure; in Greechie (1971) the author gives an example of the finite orthomodular lattice without probabilistic measures at all.

6. Thus, if μ is **a** probabilistic measure on the orthoposet $\mathcal{E}$, we explicitly define the functional $CE = CE_{\mu}$ by the following formula (where $f : I(\alpha) \to \mathbb{R}$ is **a** bet on the basis of **an** ODU $\alpha = (a(i), i \in I(\alpha))$):

$$CE(f) = \sum_{i \in I(\alpha)} \mu(a(i)) f(i).$$

For the obvious reason this functional is called the *expected utility*. In Danilov, Lambert-Mogiliansky (2008b) it is proved that is satisfies the additivity and monotonicity axioms.

7. Do other normal functionals CE different from the expected utility exist? In Danilov, Lambert-Mogiliansky (2008b) we show that they do not. Namely, we have the following

Theorem. *Let $CE : Bet(\mathcal{E}) \to \mathbb{R}$ satisfy the additivity and monotonicity. Then there exists* (**a unique**) *probabilistic measure μ on $\mathcal{E}$ such that $CE = CE_{\mu}$.*

Thus, as in the classical approach, there exists one-to-one correspondence between probabilistic measures on the orthoposet of events $\mathcal{E}$ and normal functionals of bet estimation. In particular, if there is **no** probabilistic measure, there is **no** normal CE. **Such** posets does not appear to be of natural origin. In other words, the **framework** of general orthoposets are still too wide for the profound decision-making theory. We will consider some narrowing of **the setting** in Section 8.

6 States and the updating problem

Let us return to the individual who has to choose between two (or more) bets (in the situation with the orthoposet of events $\mathcal{E}$). As we have seen, in order to do it by a universal method, he must take a probabilistic measure μ on $\mathcal{E}$. This measure can be treated as his ‘ belief’ or his ‘prior’. The latter term explicitly **hints at** appearance of a posterior, and we cannot disregard this problem.

The problem considered in Section 5 was obviously **of one-shot type**. Based on his information, knowledge, superstitions, etc., the individual (DM) forms the prior (or ‘subjective probability’) on $\mathcal{E}$ and then, using the expected utility functional, assigns ‘values’ to different bets. Real measurements (leading to one or other resolution of uncertainty and the corresponding payoff) can even be avoided. However, imagine that the measurement was performed and that the bet choice problem arises again. It is absolutely clear that after the measurement is performed, the DM’s prior can and in the general case must change. The DM has obtained new information about the state of the world, and this information can make him **to** revise his belief. How should he revise the beliefs? As previously we did not limited the DM in choosing the prior (to be more exact, the limitations were merely structural: additivity, monotonicity), so now we cannot limit him in choosing the updating either. This choice is dictated by his idea of the system, his model of the system. There is one general requirement on the updating, however.

Let us return to the classical case. Here updating occurs in accordance with the Bayes rule (or, which is the same, on the basis of conditional probabilities). Before the measurement the prior was given by a probabilistic measure β. Then the measurement was performed, which resulted in our learning that the real state of the system belongs to a subset E. How is the updated probabilistic measure β' structured then? Very simply: for the event A the new probability is $\beta'(A) = \beta(A \cap E)/\beta(E)$. In particular, $\beta'(E) = 1$.

The latter property seems reasonable in the general case as well. Let E be some (nonzero) event. By definition, there exists a measurement M and its outcome o which specify this event. No matter how the DM forms the new prior β', it is clear that $\beta'(E)$ should be equal to 1. This is because if we repeat the measurement M, we reliably obtain o. But this means that, in order that it would be possible to perform updating in any situation, our orthoposet $\mathcal{E}$ should have sufficiently many probabilistic measures. Sufficiently many in the following sense: for any nonzero event E there should be a probabilistic measure $\mu = \mu_E$ such that $\mu(E) = 1$.

This is where such an important notion as the *state* of the system inconspicuously comes into play. What is a state? Strange as it is, this question is quite difficult to answer.[3] The point is that (as in the case of probabilities) there are two approaches to this notion, subjective and objective. Is a state (in principle) merely the state of our knowledge (or our brain, if you like) or does it reflect anything objective and existing irrespective of our mind? (We willingly or unwillingly encounter the **well-**known **ghost** of the ‘main problem of philosophy.’) Or, as Penrose puts it more mildly (Penrose (1994)): is the state **no more than a convenience, useful only** for calculating probabilities for the results of measurements for the system and a representation of ‘our knowledge’ concerning the system or does the state vector actually provide **an accurate** mathematical description of **a** real quantum world? I believe that both points of view are sensible and do not contradict each other. Yes, **a** measurement first of all changes our knowledge and calls for the revision of the state. But it would be hardly necessary to change substantially our prior if anything in the system did not really change in the course of the measurement.

[3] This somewhat resembles **controversies** about the notion of probability.

'Subjectively', **a** state is our information on the system. Once the result of a new measurement is obtained, the state of the system changes merely because we (DMs) have gained new information. Note two characteristic features of this approach that will also appear in the 'objective' approach. One is that there are **sufficiently many** states; anyhow, for each outcome of the full measurement there must exist **the corresponding** state. The other is that the state determines the probabilities (subjective, of course) for any outcome of any measurement. The latter property underlies the 'objective' definition of the state. From the objective point of view, the state is something that determines the result (more exactly, probabilities of outcomes) of any measurement.

However, this most important property of states does not yet tells us what the state is and how we can describe the set S of all states. I think that this is of fundamental importance. While the set of measuring instruments is given from the very beginning, there is nothing to indicate what should be called a state. The researcher or the modeller must themselves 'guess right' what should be considered **as** a state. To guess right means to give **such** a definition that predictions of a theory or a model will agree with the results of the experiments.[4]

It seems that the notion of **a** model of the system to be introduced below is wide enough to accommodate both the subjective and the objective approach. **A** *model* consists of the following four (or five) types of data:

1. **A** set $\mathcal{M}$ of *admissible measurements* (or, somewhat longer, measuring instruments) of the first kind. For $M \in \mathcal{M}$, the set of potential outcomes of this measuring instrument is designated as $O(M)$.

2. **A** set S of *states* of the system.

3. For each state s and each measuring instrument M a *random outcome* $\mu_M(s) \in \Delta(O(M))$ should be given. (In what follows $\Delta(X)$ denotes the set of probabilistic measures on the finite set X.)

In other words, if we fixed the state s, the measurement M, and some outcome $o \in O(M)$ of this measurement, the probability $\mu_M(s)(o)$ for obtaining this outcome o of the measurement M performed in the state s should be determined. Needless to say that the sum of these figures over all $o \in O(M)$ must be 1 (some outcome must occur). These data are given in a more compact form by **a** mapping

$$\mu : S \to \times_M \Delta(O(M)).$$

Another type of data is specification of the change in the states that results from the measurements, whether it is 'reasonable' revision of priors or a 'real' alteration in the objective state resulting from the interaction of the system with the measuring instrument. To this end, we introduce a notation. Let M be a measurement and $o \in O(M)$ an outcome of this measurement. $E_M(o)$ will denote the set of those states where the outcome o occurs **with certainty** (it is natural to call it the '**eigenset** of the measurement M with the

[4] The subjective approach seems to say that **a** state is a probabilistic measure, i.e. an element of $\Delta(\mathcal{E})$. But is it any element? At first glance, yes, it is; what can restrict mental arbitrariness of the DM? Updating can! S should be the minimum set in $\Delta(\mathcal{E})$, closed with respect to updating. See also the end of Section 8.

eigenvalue o').

4. For each pair (M, o) $(o \in O(M))$ the *transition mapping* $\tau_{M,o} : S \to E_M(o)$ should be given.

The meaning of this mapping is that if the state of the system before the measurement M was s and if the result of the measurement was o, then after the measurement the state of the system is $\tau_{M,o}(s)$. That the state $\tau_{M,o}(s)$ belongs to $E_M(o)$ is just another formulation of the fact that the measurement M is of the first kind. Roughly speaking, ✓ the mapping $\tau_{M,o}$ is the projection of S onto the **eigenset** $E_M(o)$.

5. Finally, if the system possesses its own (not caused by measurements) dynamics, ✓ the model should involve the description of this dynamics, i.e., **a** semigroup family φ_t of mappings of S into S (t runs through natural numbers if time is discrete, or $\mathbb{R}_+$ if time is continuous), $\varphi_{t+s} = \varphi_t \circ \varphi_s$.

Example: the *model of hidden parameters*. Let us give a rather general principle of model construction evoked by the problem of 'hidden parameters' in quantum mechanics. As was pointed out in the Introduction, a bright and objection-arousing feature of quantum theory is its indeterminism. Opponents of the theory tried to construct models where true states were assumed to uniquely determine results of measurements while ✓ involving hidden parameters unknown to the observer. **Adherents** of the theory proved theorems of impossibility of hidden parameters; see Kholevo (2001) for an account of this intellectual war. Below a simple and tautological method for construction of a model with hidden parameters is proposed.

✓ Let $\mathcal{M}$ be a set of measurements. As the set S, we use the **disjoint union of** $O(M)$, $S = \coprod_{M \in \mathcal{M}} O(M)$. Thus, the state is the pair (M, o), where $M \in \mathcal{M}$ and $o \in O(M)$. In addition, we assume that for each state $s = (M, o)$ an element $\mu(s)$ in $\Delta(\times_M O(M))$ ✓ is given such that $\mu(s)(\{s\} \times (\times_{M' \neq M} O(M'))) = 1$. A **map** $s \mapsto \mu(o)$ gives the mapping $S \to \times_M \Delta(O(M))$ required in item 3. Finally, the mapping τ_s from item 4 is given by ✓ Here is my mistake mapping all S into the point s.

Here the set $\times_M O(M)$ is understood as the space of hidden parameters or 'true' states of the system. In each 'true' state, the outcome of any measurement M is uniquely determined. But we do not known exactly the 'new true state' after the measurement M that yielded the result o. We only know its probabilistic distribution described by ✓ $\mu(M, o)$. Note that it is exactly how the model of **a** fly in **a** box from Section 2 works.

Some additions to the general definition of the model. In item 3 we spoke about the mapping of the set of states S into the product of simplices $\Delta(O(M))$. It is quite ? natural to consider this mapping **to be** injective. The justification is as follows: we cannot distinguish states if they are indistinguishable by any measurement. But then S can be thought of as being a subset in $\times_M \Delta(O(M))$. The latter is a product of convex sets and thus it is convex itself. Though there are no reasons to hold that S coincides with this product (see the example with the fly in the box), S can be considered, without stretching ✓ a point, as a convex subset of this product.

✓ Thus, without loss of generality and sense, S may be considered **as a** convex (and ✓ closed) set. Then the subsets $E(o)$ are also convex, and the transition mappings are linear.

Convexity allows **to introduce** the notion of pure states. If the state s is the extreme point of the convex set S, it is called *pure*. This simply means that it cannot be represented as a nontrivial convex combination (or probabilistic mixture) of other states. Conversely, any other state can be represented as a mixture of pure states. Thus, instead of the set S of all states, a smaller set P of pure states can be used. The only reservation to be made refers to transition mappings. It should be remembered that in the general case a pure state can pass into a mixed state as the measurement is performed.

It is natural to call subsets like $E_M(o)$ *properties* of the system (in Section 4 we called them events); more exactly, it is the property to yield the outcome o from the measurement M. We even consider the particular case where the measurement M is dichotomic, that is, has only two outcomes. These measurements are called (following Mackey) *questions*, with one of their outcomes considered to be YES and the other NO. If a question Q is given, we get two properties: the property E_Q(YES) and the 'opposite' property E_Q(NO). We call these properties opposite because if one property is fulfilled, the other is not (i.e., these subsets in S do not intersect); in addition, in any state we get one of the answers to the question Q. But one should never think that these two subsets, being subsets in S, cover entire S. The**re** are plenty of 'intermediate' states that do not belong to either E_Q(YES) or E_Q(NO); these states are characterized by having a positive probability to give the answer YES and NO. This is just about the pre-existing properties. Let s be this 'intermediate' state (physicists use the term 'superposition', which confuses rather than clarifies the point); does the state s possess the property E_Q(YES)? No, it does not, of course. Does it possess the opposite property? No, it does not, of course. And there is nothing unusual or mysterious.

Imposing additional conditions or axioms, we can find that the set of all properties constitutes an ortholattice (atomic, orthomodular, etc.). Details can be found in Danilov, Lambert-Mogiliansky (2008a).

7 An illustration

As an illustration we consider a somewhat speculative choice problem borrowed from Danilov, Lambert-Mogiliansky (2008a). That is, it is not an example from 'real life' (experiments) but a fantasy about what there could be.

Imagine that a DM is given various problems of choosing between three alternatives a, b, c (e.g., apple, banana, lemon). Thus, here **a** measurement is **a** choice. There are four different measurements (or choice problems). The first measurement-choice is that the person is offered to choose (one and only one) object from the set $\{a, b, c\}$. The outcomes of this measurement are a, b and c; for brevity, we denote the measurement itself by abc. The other three measurements correspond to the pairs of alternatives. For example, the measurement ab is the choice of the object from the pair $\{a, b\}$. Other pairs are treated in a similar way. It is assumed that all four measurements are of the first kind. This completes the description of data 1 from the definition of the model.

We also impose the following assumption of *rationality* (or **compatibility**) of two neighboring measurements. Namely, imagine that the measurement abc yielded the result x. If immediately after that we perform the measurement xy ($x \neq y$), its outcome must be x. And vice versa, let us perform the binary measurement xy and obtain the result x. Then, if the next measurement is abc, its outcome CANNOT be y.

We next consider three models possible for this system.

Model 0. Let us assume that all measurements-choices are consistent and therefore we can perform them in any order without changing results of measurements. The resulting compound measurement has, as is readily seen, six outcomes. For example, if the measurement abc yields the outcome a, the outcomes of the measurements ab and ac are obvious from the rationality condition, and it only remains to perform the measurement bc, which yields two outcomes. If the outcome obtained from the measurement bc is b, we can treat everything as if the individual had the preference $a > b > c$; other outcomes are interpreted in a similar way. And we can define the set of pure states as the set consisting of six linear orders on $\{a, b, c\}$. Measurements do not alter pure states. In a word, we get the classical theory of rational choice.

Let us now consider a more general case where measurements are not assumed to be consistent.

Model 1. Symbols $[a]$, $[b]$, and $[c]$ are claimed to be pure states. The effect of measurements on the states and the transition probabilities are specified below:

a) If we perform the measurement abc, the system in the state $[x]$ does not change the state and yields the outcome x;

b) If we perform the measurement ab, the state $[a]$ does not change and a is chosen, the state $[b]$ does not change either and b is chosen. But if the state of the system is $[c]$, it yields the outcome a or b with equal chances and undergoes a transition to the state $[a]$ or $[b]$ respectively.

The similar holds for the measurements ac and bc.

Model 2. However, another model can be proposed for the same system. Now it will involve nine pure states. Roughly speaking, these states correspond to nine outcomes of our four measurements. Three states are designated as $[a]$, $[b]$, and $[c]$; they correspond to the measurement abc. The measurement ab has two outcomes; accordingly, we introduce two states designated as $[a > b]$ and $[b > c]$. The states $[a > c]$, $[c > a]$, $[b > c]$, and $[c > b]$ are introduced in a similar manner.

Now we should describe probabilities for outcomes and the corresponding state transitions.

a) Assume that we perform the measurement abc. As far as the states $[a]$, $[b]$, and $[c]$ are concerned, everything is obvious: they do not change and the corresponding result is obtained. It is more interesting what takes place in the state $[a > b]$. The result of the measurement is a with probability 2/3, and the system changes its state to $[a]$; the measurement result is c with probability 1/3, and the system changes the state to $[c]$; b cannot occur. The situation is similar (symmetrical) for other states.

b) Assume that we perform the measurement ab. In the state $[a]$ the outcome is a and the state changes to $[a > b]$. In the state $[b]$ the outcome is b and the new state $[b > a]$ arises. In the state $[c]$ the outcome a (and the transition to the state $[a > b]$) and the outcome b (and the transition to the state $[b > a]$) have equal chance of occurring. In the state $[a > b]$ the outcome of the measurement ab is a and the state does not change; in the state $[b > a]$ it is vice versa. In the state $[a > c]$ the result of the choice is a with probability 2/3 (and the system undergoes transition to $[a > b]$) and b with probability 1/3 (and the system undergoes transition to $[b > a]$). In the state $[c > a]$ it is vice versa: a is chosen with probability 1/3 and b with probability 2/3.

For the other two measurements ac and bc everything is symmetrical.

It is easy to understand that the rationality condition (**compatibility** of two neighboring measurements) in model 2 is also fulfilled.

Above we have proposed two models describing the nonclassical rational choice. There are certainly a great deal of other models differing by both probabilities and states. Which of them is **true** (or more **true**) can only be shown by comparing their predictions and the observations. These models predict quite different behavior. Model 1 describes a totally inert chooser; having chosen an object, he persists making this choice as long as he has a possibility of doing it. Only if this subject is not on the list, he randomly chooses any of the available ones (and again keeps showing loyalty to it). The behavior predicted by **the** model 2 is more interesting.

Anyway, both models agree with the above citation from Busemeyer, Wang (2007). Let us consider for example **the** model 2. Assume that we have offered making a choice from abc and obtained a. If we repeat the offer, we will have a again. Now imagine that we inserted a filler ab between two measurements-choices abc. Its outcome is obvious, but we have a transition to the state $[a > b]$. When we offer abc again, we will have the result c with probability 1/3. Thus, model 2 predicts quite definite probabilities which one can test and either agree or disagree with the model.

A similar example. Assume that a person has chosen a from ab. If the problem is immediately repeated, he (according to **the** model 2) will choose a again. But we insert a filler problem abc in between the repetitions. Then the new state will be a mixture $2/3[a] + 1/3[c]$. We repeat the measurement ab. In the state $[a]$ the choice will be a, and in the state $[c]$ the choices a and b will have equal chance of being made. Therefore, the repeated measurement ab yields a with probability 5/6 and b with probability 1/6. Thus, the reversal of the initial preference $a > b$ occurs with probability 1/6.

8 Transition probability spaces

In this section we briefly discuss a rather general method for construction of models proposed in Mielnik (1968) and elaborated in Pulmanova (1986). In this method we begin with states and transition probabilities rather than with measurements. The transition probabilities make it possible to determine both measurements and changes in states resulting from the measurements.

Definition. **A** *transition probability space* is **a** pair (S, τ), where S is a set and τ is the mapping from $S \times S$ into the segment $[0, 1]$.

These data should satisfy three axioms T1–T3 which we formulate below. But it is better to begin with the interpretation: elements S are regarded as 'pure' states, and the number $\tau(s, t)$ for $s, t \in S$ is regarded as the probability for the transition from s to t under the effect of an appropriate measurement (or as a measure of proximity of the states s and t).

Axiom T1. If $\tau(s,t) = 0$, then $\tau(t,s) = 0$.

If $\tau(s,t) = 0$, we say that the states s and t are *orthogonal*, $s \perp t$. By virtue of T1 the orthogonality relation $\perp$ is symmetrical (the function τ itself is not considered symmetrical). From axiom T2 that follows, it will be clear that the relation $\perp$ is also irreflexive. So the pair $(S, \perp)$ is an *orthospace* in the sense meant in Danilov, Lambert-Mogiliansky (2008a). In particular, we will denote **by** $A^{\perp}$ **the** set orthogonal to **a** subset $A \subset S$, i.e.

$$A^{\perp} = \{s \in S,\ s \perp a \text{ for any } a \in A\}.$$

Axiom T2. $\tau(s,t) = 1$ if and only if $s = t$.

In particular, $\tau(s,s) = 1$, and thus no state is orthogonal **to** itself. It is actually **the** first-kindness again. The second part of this axiom says that if under the appropriate measurement the state s always passes into t, then s and t coincide.

We have already used the term 'appropriate measurement' several times. Now it is time to clear it up. First, there is a couple of notions. **A** subset $B \in S$ is called *orthogonal* if $a \perp b$ for **different** a, b from B. **An** orthogonal subset maximal with respect to inclusion is called **an** *orthobasis* in S. Orthobases are regarded as full measurements. To make it reasonable, the third and last axiom is imposed:

Axiom T3. $\sum_{b \in B} \tau(s,b) = 1$ for any orthobasis B and any $s \in S$.

Let us return to the interpretation of **an** orthobasis B as **a** (full) measurement. Outcomes of this measurement are identified with elements B. If the system is in **a** state s and if the measurement B is performed, the system passes into **a** state b with probability $\tau(s,b)$ and outputs the corresponding signal b. Axiom T3 ensures that the set of these numbers can really be understood as the probability. In particular, if the initial state s itself belongs to B, it remains in place (according to T2).

What is good about transition probability spaces is that axioms M1–M3 from Section 4 can be justified for them. And events themselves can be represented as special subsets in S. Recall that in Section 4 events were constructed on the basis of the relation $\Longrightarrow$ on the set of pre-events. In currently used terms the pre-event is an arbitrary orthogonal set. Let A be an orthogonal set in S, **and let** $E(A) = \{x \in S,\ \sum_{a \in A} \tau(x,a) = 1\}$. In terms of content, $E(A)$ consists of states x in which the outcome of **an** appropriate full measurement falls within A **with certainty** (recall the definition of **an eigenset** from Section 6). So let orthogonal sets A and A' be regarded as pre-events. Then the relation $A \Longrightarrow A'$ is equivalent to the fact that $A \subset E(A')$. The following statement holds (see Danilov, Lambert-Mogiliansky (2008b)):

Lemma. *For* **an** *orthogonal subset A the equality $E(A) = A^{\perp\perp}$ is valid.*

Now it is relatively easy to see that properties M1 and M2 are satisfied, that the event additional to A exactly equals $A^{\perp}$, and that the set $\mathcal{E}$ of events is identified with the set of orthoclosed subsets in $(S, \perp)$ which have **an** orthobasis. In addition, it can be shown that the orthoposet $\mathcal{E}$ is orthomodular.

We also argue that elements of the set S can be regarded as probabilistic measures on the orthoposet $\mathcal{E} = \mathcal{E}(S, \perp)$. Let $E = E(A)$ be **an** event, where A is **an** orthogonal subset. For the element $s \in S$ we put $s(E) = \tau(s, A) := \sum_{a \in A} \tau(s,a)$. (Intuitively, this is

✓ the probability that **an** appropriate measurement will yield the outcome in A.) It can be shown that this definition is correct, i.e., $s(E)$ depends only on E and not on A. And it it almost obvious that s is the probabilistic measure on $\mathcal{E}$. Thus, we obtain the canonical mapping

$$S \to \Delta(\mathcal{E}).$$

It can be verified (see Danilov, Lambert-Mogiliansky (2008b)) that this mapping is injec- ✓ tive and that elements of S are extreme points of S as the subset in $\Delta(\mathcal{E})$ (which justifies their initial interpretation as pure states).

✓ It is pertinent to return to the question of states raised in Section 6. We have said that the elements of $\Delta(\mathcal{E})$ can be regarded as a priori beliefs. But the elements of S (or its convex hull co(S) in $\Delta(\mathcal{E})$) can already be treated as real states (mixed, in the case ✓ of co(S)). This is because we not **only** are able to calculate probabilities for outcomes of any measurement but know how states change under the effect of measurements.

Thus, the transition probability spaces automatically yield models of measured systems. It is quite a flexible and convenient (though not universal) tool for construction of models. Let us give a few examples.

Example 1: classical. Let S be an arbitrary set **with trivial transition probabil-** ✓ **ities**: $\tau(s,t) = 1$ if $s = t$, and $= 0$ in other cases. There is the unique full measurement, ✓ (pure) states remain unchanged, **an** event is an arbitrary subset in S, and **a mixed state** ✓ **is an arbitrary probabilistic measure on** S.

In general, if (S, τ) is an arbitrary transition probability space, S can be decomposed into 'connected' pieces. Two elements s and t belong to one connected component if they can be connected by a chain of nonorthogonal states. As a result, we obtain decomposition of S to an 'orthogonal sum' of connected components S_j, $S = \oplus_j S_j$. These connected components S_j can be regarded as *superstates* or as the 'classical' factor of the system. ✓ Inside each superstate we have a **genuine** nonclassical system.

✓ It is '**genuine** nonclassical' in the sense that it is already irreducible, it cannot be decomposed into smaller subsystems. A transition from any state of the irreducible system to any other state can be carried out using an appropriate sequence of measurements. An ✓ irreducible system is a 'thing in itself'; revealing **some its side** it remains unpredictable on the other sides. It is a 'point' or an 'atom' which has intrinsic degrees of freedom. The examples given below are irreducible systems.

Example 2: the Hilbert space model. Let H be a finite-dimensional Hilbert space (over real numbers, for simplicity; i.e., simply an Euclidian space). As S, we use a set of one-dimensional linear subspaces ('straight-line') in H. If L and L' are two straight lines, then $\tau(L, L') = \cos^2(\varphi)$, where φ is the angle between these straight lines. (If L and L' ✓ are given by unit vectors v and v', the transition probability is the square of the scalar product of v and v'; it is the famous Born rule.) Axioms T1 and T2 are obvious while T3 is Parseval's equality (or the Pythagorean theorem). Gleason's famous theorem states ? that (**under** the assumption that dim $H \geq 3$) $\Delta(\mathcal{E}) = co(S)$.

This is exactly the standard *Hilbert space model* of quantum mechanics. General ✓ irreducible transition probability spaces can be regarded as **a** non-Abelian version of Hilbert spaces.

Example 3: a fly in the box. Here the set S consists of four pure states $\{L, R, F, B\}$. The transition probabilities $\tau(s,t)$ are given by the following table:

$s \diagdown t$	L	R	F	B
L	1	0	p	$1-p$
R	0	1	p'	$1-p'$
F	q	$1-q$	1	0
B	q'	$1-q'$	0	1

Finally, note that model 2 of preferences from Section 7 does not fit into transition probabilities.

Conclusion

Does the aforesaid have to do with the systems studied by biological or social sciences? In other words, do inconsistent measurements occur in them? We would like to believe they do. But the question should most likely be formulated in a different way: do measurements of the first kind (say, those not reduced to physical measurements like weighing) occur in them? **Only for them we can discuss their** inconsistency. If the measurement is not of the first kind, it is already inconsistent with itself. It is no longer a measurement but **a** fortune-telling.

Assume that inconsistent measurements do exist and thus the macrosystem shows quantum-like behavior. Here are some considerations as to why this system could change its state upon undergoing a measurement.

A measurement makes the system to respond to a question or a challenge to which it (system) is generally not prepared. (If an answer is already available, it is simply extracted from the depth of the memory, and the state does not change.) The system looks for the answer to the question asked, it uses some of its resources to **build** a structure which allows the question to be answered. This auxiliary structure is being **built** quite randomly. The system makes a try, a second, a third, and here it is at last, **a** solution to the problem! The answer can be given. (Imagine that you are caught by hail, you are looking for a safe shelter, running about, and in the end find (one of) such shelters.) If this measurement-trial (in the sense of "**torture**") will be immediately repeated, the system need not strain itself for the solution is already found. Therefore, the answer is just repeated, and that is whence first-kindness arises.

Imagine however that the system is faced with a new test and has to find the answer to another (somewhat similar yet different) question. Different in the sense that the previously built structure does not allow an answer to the new question, and similar in the sense that the same resources as were used before should be employed to find the answer to it. Therefore, solution of the new challenge problem calls for destroying the previously built system and building a new one from its remains. (Recall the example with the memory cell from Section 2.) The answer is again undetermined, random. But the important point here is that if after getting the answer to the second question we ask the first question again, the system will have to destroy the second structure and to build the answer to the first question from its elements. (It is like a meccano. You are asked to make a car. And then a plane, from the same parts. And a car again.) And it is not sure that the previous structure will be reconstructed: the search is carried out for the most part blindly, at random.

References

Mackey G. (1963) *Mathematical Foundations of Quantum Mechanics.* Benjamen, New York.

Penrose P. (1994) *Shadows of the Mind.* Oxford University Press, Oxford.

Feynman R., Leighton R., and Sands M. (1965) *The Feynman Lectures on Physics*, vol. 3, Addison-Wesley, London.

von Neumann I. (1932) *Mathematische Grundlagen der Quantemmechanik.* Springer-Verlag, Berlin.

Kholevo A.S. (2001) *Statistical Structure of Quantum Theory.* Springer-Verlag, Berlin.

Beltrametti E.G., Cassinelli J. (1981) *The logic of quantum mechanics*, Encyclopedia of Mathematics and its Applications, Ed. G.-C. Rota, Vol. 15, London, Addison-Wesley.

Birkhoff G., von Neumann J. (1936) The logic of quantum mechanics, *Ann. Math.*, 37, 823-843.

Busemeyer J.R. and Z. Wang (2007) Quantum Information Processing Explanation for Interactions between Inferences and Decisions. In *Quantum Interaction* (Papers from the 2007 AAAI Spring Symposium), AAAI Press, 91-97.

Coecke B., Moore D., Wilce A. (2001) Operational quantum logic: an overview. In *Current Research in Operational Quantum Logic: Algebras, Categories, Languages* (eds. Coecke B., Moore D., and Wilce A.), Kluwer Academic Publi shers, 2001. Also available at arXive: quant-ph/0008019.

Danilov V.I. and A. Lambert-Mogiliansky (2008a) Measurable systems and behavioral sciences, *Mathematical Social Sciences* 55, 315-340.

Danilov V.I. and A. Lambert-Mogiliansky (2008b) Decision-Making under Non-classical Uncertanty. In 'Quantum Interaction (2008)', 83–86.

Franco R. (2008) The inferse fallacy and quantum formalism. In *Quantum Interaction* (2008) , 94-97. Available on http://xxx.lanl.gov/abs/0708.2972.

Greechie R.J. (1971) Orthomadular lattices admitting no states. *Journal of Combinatorial Theory* 10, 119-132.

Holland S.S. (1995) Orthomodularity in infinite dimensions: a theorem of M. Solèr. *Bull. of AMS*, 32, 205-234.

Kahneman D. and Tversky A. (2000) *Choice, Values and Frames.* Cambridge, Cambridge Universtity Press.

Lambert-Mogiliansky A., S. Zamir, and H. Zwirn (2004) Type-indeterminacy - A Model for the KT-(Kahneman and Tversky)-man. xxx.lanl.gov/physics/0604166

Mielnik B. (1968) Geometry of Quantum States. *Commun. Math. Phys.* 9, 55-80.

Pulmanova S. (1986) Transition probability spaces, *Journal of Mathematical Physics* 27, 1791-1795.

Quantum Interaction (2008) Proceedings of the Second Quantum Interaction Symposium (eds. P.Bruza, W.Lawless, K. van Rijsberger, D.Sofge, B.Coecke, S.Clark) London, 2008.

Quantum Theory and Measurement (1983) (eds. J.A. Wheller and W.H. Zurek), Princeton, Princeton Univ. Press.

Savage L. (1954) *The Foundations of Statistics.* New York, John Wiley.

Tversky A. and E. Shafir (1992) The disjunction effect in choiceunder uncertainty. *Psucological Science* 3/5, 305-309.

Wright R. (1978) The state of the Pentagon, in *Mathematical Foundation of Quantum Theory* (ed. A. R. Marlow), Academic Press New York San Fransisco London.

Wright R. (2003) Statistical structures underlying quantum mechanics and social science. arXiv:quant-ph/0307234.