A property type invariant model of measurement applied to nominal evaluations

2015 J. Phys.: Conf. Ser. 588 012008

(http://iopscience.iop.org/1742-6596/588/1/012008)

A property type invariant model of measurement applied to nominal evaluations

Luca Mari[1,#] and Arianna Mencattini[2]

[1] Università Cattaneo – LIUC, School of Industrial Engineering, Italy
[2] University of Rome Tor Vergata, Dept. of Electronic Engineering, Italy

lmari@liuc.it

Abstract. The scientific community and the standardization world are exploring the concepts of nominal evaluation and nominal property in the metrological context. In this paper we show that measurement can be framed in a structural context that is independent of the algebraic structure assumed for the measurand and the measured values. In such a framework the basic metrological concepts are consistently applied also to the nominal case.

1. Introduction: the problem of measurability of nominal properties

Measurement is loaded of stereotypes, and some of them are related to the concept itself of measurability: *what can be measured?* Indeed, expressions such as "counting and measuring" and "weights and measures" suppose that counting and weighing are not ways to measure. It is the plausible trace of the Euclidean standpoint, in which only geometric quantities were considered object of mensuration, intended as a mathematical activity aimed at comparing quantities by their ratio. The theory underlying Lebesgue measure is the outcome, where not confusing "measure" and "measurement" here is then crucial [1]. Referring to this tradition to justify that only ratio (or even only additive) quantities are measurable is then a weak strategy. It is true that the variables appearing in physical laws generally denote ratio quantities, and that the powerful tools of quantity calculus / dimensional analysis apply only to ratio quantities, that are structurally more informative than, e.g., purely ordinal properties. But since values of ratio quantities can be obtained by means other than measurement, e.g., subjective estimation and guess, being a ratio quantity is not sufficient for being object of measurement: is it necessary? [2] Not so surprisingly, a negative answer comes from the *International Vocabulary of Metrology* (VIM) [3], that in its latest edition extends measurement to ordinal entities and makes the move lexically explicit by terming them "ordinal quantities", an expression that would be considered contradictory in the Euclidean framework. The VIM also defines 'nominal property', at the same time stating that "measurement does not apply to nominal properties" (and in fact the chosen term is not "nominal quantity"), because they "have no magnitude", thus implicitly implying that "having magnitude" is a necessary condition for a property to be measurable (note that 'magnitude' is not defined by the VIM; is it, in particular, identical to 'rank', as in "shape cannot be measured because objects cannot be ranked with respect to shape" [4] p.99 ?). Since measurement is a designed-on-purpose, instead of natural, process, one is surely free to characterize it as she considers the best, and in fact multiple strategies of definition have been followed [5]. What

[#] One of the authors is a member of the Joint Committee on Guides in Metrology (JCGM) Working Group 2 (VIM). The opinion expressed in this paper does not necessarily represent the view of this Working Group.

does not seem so conventional is the good quality assumed of the measurement results (intended as accuracy, reliability, ...), that is not generally expected from, e.g., subjective estimations and guesses. On the other hand, this criterion is in principle unrelated to the condition of having magnitude, so that activities are in progress (both in the scientific community, e.g., [6], and in the standardization world, e.g., [7]) to better explore the concepts of nominal property and its evaluation, in a metrological context and with a metrological approach.

Our contribution here is to show that *the invariant features of a measurement process are independent of the algebraic structure of the measurand*, and therefore that from the point of view of the structure of the process the concept of measurement could be safely generalized so to encompass nominal properties. The analysis will lead us to show, in particular, that the usual contraposition measurement vs classification is misleading, as classification is just a component of the process of "reasonable attribution of one or more property values to a nominal property" (a paraphrase of the VIM definition of 'measurement', where "quantity" is used instead of "nominal property"). Whether this generalization will be lexically accepted is mainly a matter of social conventions.

2. Nominal properties and nominal evaluations

In the context of measurement science nominal properties have been often treated in confused way. Consider for example: "The usual method of assigning code numbers to entities or unordered classes is, in effect, to select integers randomly (without replacement) from the sequence 1, 2, ..., k. It is because numbers so selected bear no relationship to the 'relative amount or degree of a property possessed by the object' or class that the construction of such 'nominal scales' is not measurement." [8] p.127. Together with the peculiar hypothesis that a scale construction is a measurement, the implicit message is conveyed here that nominal properties are not empirical entities, and therefore assigning values to them is a matter of conventions (as in the unfortunate example given by the VIM: "ISO two-letter country code") or even randomness. In what follows we show that this is just wrong.

Two quantities, and more generally two properties of the same kind are by definition mutually comparable. The simplest condition of comparability is indistinguishability: if a property P_a of an object a and a property P_b of an object b are of the same kind P then the two properties are either indistinguishable, $P_a = P_b$, or $P_a \neq P_b$ ($P_a = P_b$ can be considered conveying the same information as $a \approx_P b$, i.e., the objects a and b are indistinguishable with respect to the general property / kind P, "P-indistinguishable" for short). While for some properties this information can be refined, for example by splitting the case $P_a \neq P_b$ as either $P_a < P_b$ or $P_a > P_b$, for some other properties the only comparison that conveys meaningful information is the one of P-indistinguishability. The entities for which this happens are called *nominal properties*.

By now recalling the VIM definition of 'measurement', "process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity", the basic issue arises: is it possible, and in the case how, to design, setup and perform *a process of experimentally obtaining one or more property values that can reasonably be attributed to a nominal property*, in such a way that the same "reasonability" criterion applies in both cases?

We believe that the process of "reasonable attribution" of a value to the property P_x, for an object x under consideration, can be intended in measurement science in a purely structural way, as follows.

S1. *Define a set of reference properties*: a set P^* of properties P_i^* of the same kind P is defined, such that if $i \neq j$ then $P_i^* \neq P_j^*$ for all $P_i^*, P_j^* \in P^*$.

S2. *Establish a criterion of P-indistinguishability with reference properties*: a criterion is defined to operatively assess that for any candidate property P_x there is one and only one reference property $P_i^* \in P^*$ such that $P_x = P_i^*$. This may require producing a set S of objects s_i, each of them realizing a reference property P_i^*, so that the assessment is performed by comparing objects, $x \approx_P s_i$, or it could be structured as an algorithm that recognizes the patterns P_i^*.

S3. *Attribute a value to P_x*: the defined criterion is applied, for example by comparing the candidate object x with the objects in S with respect to the general property P, so that once the object s_i is found such that $P_x = P_i^*$, P_i^* as value of P_x is attributed.

Some notes on this process.

First and most important. The steps S1-S3 are independent of the type of the involved properties, and thus apply indifferently to both, e.g., nominal properties and ratio quantities. The fact that in the latter case some specific techniques can be adopted (see the next note) does not eliminate this commonality.

Second. The step S1 is sometimes considered of "scale construction". In the case of ratio quantities, for which there is a "natural zero" property, it can be performed by choosing a "unit" property, P_1^*, and then obtaining P_i^* by concatenating P_1^* i times with itself. A nominal property does not have a "natural zero" nor a concatenation, and therefore this construction procedure cannot be exploited and the reference properties in P^* must be explicitly listed (note that the same applies to ordinal quantities).

Third. The step S2 usually includes the process of "standards dissemination": the objects with which the candidate object is possibly compared are not those originally produced (the "primary standards") but some suitably calibrated replicas of them, so to guarantee the (metrological) traceability to P^*. This also shows that a critical requirement is the stability of the reference properties in P^*.

Fourth. While the generic linguistic form of property values is "P_i^* in P^*" (as in "3.5 in Richter scale"), in the case of ratio quantities the name of the unit, "P_1^*", can be considered as the name of the whole reference set, and property values can be denoted as "i P_1^*" (as in "3.5 m"), actually meaning "i times P_1^*".

Fifth. Let us call *nominal evaluation* the process S1-S3 when applied to a nominal property. Then nominal evaluations and classifications (intended as clustering of objects according to property) are distinct processes, and while classification is induced by nominal evaluation (two objects belong to the same class if there are evaluated as P-indistinguishable) the converse does not hold (consider, e.g., that a classification might be obtained without the previous choice of reference properties, and therefore with no constraints on traceability).

Sixth and as a synthesis. The good quality assumed of the measurement results is obtained by suitably designing, setting up and performing the process S1-S3, in particular:
- by a properly refined criterion of P-indistinguishability, so that the set P^* of reference properties includes many elements (and quantity values have a high number of significant digits in the case of ratio quantities);
- by selecting stable and universal reference properties (such as the speed of light in vacuum or the electron charge);
- by producing stable primary and working standards, thus building an appropriate traceability chain; by experimentally performing the comparison in S3 by means of high quality instrumentation, thus guaranteeing the objectivity of the outcome.

Interestingly, all these conditions are independent of the type of the involved properties.

3. A note on multivariate nominal evaluation

Exactly as quantities, also nominal properties can be multivariate, but the lack of algebraic structure allows us to avoid the peculiar consideration of the VIM: "A quantity as defined here is a scalar. However, a vector or a tensor, the components of which are quantities, is also considered to be a quantity." [3] 1.1 Note 5. The simpler characterization is here that a nominal property can be multivariate, and nevertheless the structure of the process S1-S3 is maintained.

This can be easily shown. Let $P_{(k)i}^*$ be the i-th reference property for the k-th component of the multivariate general property $\langle P_{(k)} \rangle$ under consideration. Then:

S1. *Define a set of reference properties for each $P_{(k)}$ in $\langle P_{(k)} \rangle$*: a set $P_{(k)}^*$ of properties $P_{(k)i}^*$ of the same kind $P_{(k)}$ is defined, such that if $i \neq j$ then $P_{(k)i}^* \neq P_{(k)j}^*$ for all $P_{(k)i}^*, P_{(k)j}^* \in P_{(k)}^*$.

S2. *Establish a criterion of $P_{(k)}$-indistinguishability with reference properties for each $P_{(k)}$ in $\langle P_{(k)} \rangle$*: a criterion is defined to operatively assess that for any multivariate candidate property $\langle P_{(k)x} \rangle$ there is one and only one reference property $P_{(k)i}^* \in P_{(k)}^*$ such that $P_{(k)x} = P_{(k)i}^*$.

S3. *Attribute a value to* $\langle P_{(k)x} \rangle$: the defined criterion is applied for each $P_{(k)}$ in $\langle P_{(k)} \rangle$ so to find $P_{(k)x} = P_{(k)i}^*$, and on this basis $\langle P_{(k)i}^* \rangle$ as value of $\langle P_{(k)x} \rangle$ is attributed.

This formulation frames nominal evaluations in a very general context, in which the multivariate properties $\langle P_{(k)x} \rangle$ are classified by their evaluation, and marginal (and less refined) classifications are obtained by clustering according to a partial criterion of $P_{(k)}$-indistinguishability that applies only to some components of $\langle P_{(k)} \rangle$.

On this basis, in the following section we provide a practical example of nominal evaluations.

4. An example of nominal evaluations in clinical practice

The perception of breast tumors may be difficult, especially for small lesions. To assist radiologists in standardly reporting the detected breast lesions, the American College of Radiology (ACR) occasionally publishes or updates the Breast Imaging-Reporting and Data System (BI-RADS) [9], used by medical professionals to communicate a patient's risk of developing breast cancer and the characteristics of the abnormal signs found in the mammogram. According to the BI-RADS, a breast lesion should be described in terms of its margin, shape, and density. Hence, using the evaluation process proposed above, it is assumed that a lesion is characterized by a bivariate property $\langle P_{(1)}$ = margin, $P_{(2)}$ = shape$\rangle$, whose components are respectively the margin of the lesion and the shape of the region delimited by the margin. Steps S1-S3 can be then applied to perform an evaluation.

S1. *Define a set of reference properties for* $P_{(1)}$ and $P_{(2)}$: in this case the sets of reference properties are described as follows (the meaning of the adjectives by which the properties are described is not explained here for the sake of brevity):

$P_{(1)1}^*$ = Circumscribed, $P_{(1)2}^*$ = Microlobulated, $P_{(1)3}^*$ = Indistinct, $P_{(1)4}^*$ = Obscured, $P_{(1)5}^*$ = Spiculated;

$P_{(2)1}^*$ = Round, $P_{(2)2}^*$ = Oval, $P_{(2)3}^*$ = Lobular, $P_{(2)4}^*$ = Irregular.

From these adjectives the reader can easily, and correctly, infer that each component is nominal, as in turn the bivariate property is.

S2. *Establish a criterion of* $P_{(k)}$*-indistinguishability with reference properties for* $P_{(1)}$ and $P_{(2)}$: in this case this requires establishing a criterion that, given a lesion x, is able to recognize the i-th reference property $P_{(1)i}^*$ from which its margin $P_{(1)x}$ is indistinguishable and the j-th reference property $P_{(2)i}^*$ from which its shape $P_{(2)x}$ is indistinguishable. To this goal a computer-assisted analysis system is usually exploited, that performs the following steps:

i) automatically detect the lesion margin and the region delimited by the margin [10];
ii) compute a set of numerical features $\{f_M\}$ on the margin such as fractal features [11];
iii) compute a set of numerical features $\{f_S\}$ on the shape of the region delimited by the margin, such as eccentricity, compactness, circularity, elongation [12];
iv) apply a supervised machine learning algorithm to the extracted features $\{f_M , f_S\}$ in order to assign margin and shape values among those identified in step S1.

S3: *Attribute a value to* $\langle P_{(1)x}, P_{(2)x} \rangle$: in this case this requires applying the computer-assisted analysis system setup in step S2 to a lesion x identified in a mammogram and assign a pair $\langle P_{(1)i}^*, P_{(2)j}^* \rangle$ to the bivariate property margin-and-shape, MS for short.

On this basis, what follows aims at showing two facts. First, that systems based on nominal evaluations can convey high quality information and lead to high performance decisions. In particular, margin and shape may allow to correctly recognize the histological characteristic of a large amount of lesions. Second, that results of nominal evaluations may be further elaborated by means of "indirect evaluation" procedures to compute further properties, mass *malignancy* in this case.

In order to provide a demonstrative example of these facts, using statistically well-founded data, let us consider the largest public dataset of reported mammographic images available for the community: the Digital Database for Screening Mammography (DDSM) [13]. We collected the margin and shape assessed by expert radiologists from different hospitals for the 1714 lesions in the DDSM (828 malignant tumors and 886 benign masses), and we implemented, as an indirect evaluation, a decision tree classifier to automatically assign malignancy/benignity to each lesion, using a leave-one-patient-out cross-validation strategy. To show the first fact, we implemented a tree-based classifier trained on a single component, i.e., one for margin and one for shape. Performance is evaluated in terms of the proportion of actual results identified by each algorithm (malignant tumors and benign masses correctly classified by the algorithm). To prove the second fact, we implemented a tree-based classifier trained on the bivariate property, margin and shape, and computed the same performance metrics as before.

The table below lists the values for the bivariate classifier and those obtained for each of the univariate classifier. It can be noticed, considering the bolded values highlighted in the table, that nominal evaluation of margin alone or of shape alone can lead to good results in terms of the recognition of true malignant tumors and true benign masses respectively. However, when the two targets should be simultaneously achieved (i.e., recognition of malignant tumors as well as of benign masses), then nominal evaluation based on the bivariate property (margin and shape) may increase the overall proportion of correctly identified results. This solution simultaneously reduces the recall rate (a patient with a lesion erroneously recognized as malignant is recalled for biopsy) or the missing rate (a patient with a tumor erroneously recognized as benign is not recalled as it should be with a consequent cancer spread out in the breast).

	Margin-Shape classification tree	Margin classification tree	Shape classification tree
Proportion of correctly recognized results	0.78	0.75	0.73
Proportion of malignant tumors recognized	0.74	0.84	**0.59**
Proportion of benign masses recognized	0.81	**0.67**	0.86

5. Conclusions

In other papers, e.g., [14], we proposed that measurement should be characterized in pragmatic terms, as a property evaluation process structured in such a way to be *able to guarantee, in principle, the objectivity and intersubjectivity of its results*, where then the provided information should be:
- specific to the measurand and independent of any other property of the object or the surrounding environment: hence objective;
- interpretable in the same way by different users in different places and times: hence intersubjective.

In order to comply with the expectations of objectivity and intersubjectivity, measurement science and technology refined their means in the course of time, but their strategy remained basically the same, being based in particular on:
- measuring systems that are stable and selective, i.e., that maintain their metrological features in the inter-calibration time intervals and minimize the effects of influence quantities on produced results;
- traceability chain that allow refer measurement results to universally accessible units.

In the perspective explored in this paper, the steps S1-S3 can be thought as the structural low-level implementation of these features / requirements. The fact that these steps are independent of the property types shows that nothing prevents to design, setup, and operate processes of nominal evaluation that, in the sense mentioned above, lead to objective and intersubjective results. Of course, the differences in reference set / scale construction and invariant transformations (and therefore, critically, methods for uncertainty evaluation) between, e.g., nominal properties, ordinal quantities, and ratio quantities cannot be neglected [15], but whether nominal evaluation should be included in the context of measurement or maintained in a parallel realm seems to be just unrelated to criteria of experimental quality of results: the public trust that distinctively characterizes measurement can be given also to measurement-like (i.e., structured according to S1-S3) nominal evaluations.

References

[1] Bunge M 1973 On confusing 'Measure' with 'Measurement' in the methodology of behavioral science *The methodological unity of science* ed Bunge M (Dordrecht-Holland: D. Reidel)
[2] Mari L, Maul A, Torres Irribarra D and Wilson M 2013 Quantification is neither necessary nor sufficient for measurement *IOP J Phys Conf Ser* **459 012007**
[3] JGCM 200:2012 *International Vocabulary of Metrology – Basic and general concepts and associated terms* (VIM, 2008 edition with minor corrections)
[4] Savage CW 1970 *The measurement of sensation: a critique of perceptual psychophysics* (Berkeley: University of California Press)
[5] Mari L 2013 A quest for the definition of measurement *Measurement* **46** 2889-2895
[6] Watanabe H 2005 Coarse-grained information in formal theory of measurement *Measurement* **38** 295–302
[7] Nordin G, Dybkaer R, Forsum U, Fuentes-Arderiu X, Schadow G and Pontet F 2010 An outline for a vocabulary of nominal properties and examinations – Basic and general concepts and associated terms *Clin Chem Lab Med* **48(11)** 1553–1566
[8] Duncan OD 1984 *Notes on social measurement. Historical and critical* (New York: Russell Sage Foundation)
[9] American College of Radiology (ACR) 2003 *Breast Imaging Reporting and Data System Atlas (BI-RADS® Atlas)* (Reston, Va: American College of Radiology)
[10] Mencattini A, Salmeri M, Casti P, Raguso G, L'Abbate S, Chieppa L, Ancona A, Mangieri F and Pepe ML 2011 Automatic breast masses boundary extraction in digital mammography using spatial fuzzy c-means clustering and active contour models *IEEE Int. Workshop on Medical Measurements and Applications (MeMeA)*
[11] Rangayyan RM and Nguyen TM 2007 Fractal analysis of contours of breast masses in mammograms *J. Digital Imaging* **20(3)** 223–237
[12] Mencattini A, Salmeri M, Rabottino G and Salicone S 2010 Metrological characterization of a CADx system for the classification of breast masses in mammograms *IEEE Trans Instr Meas* **59(11)** 2792–2799
[13] Heath M, Bowyer KW, Kopans D, Moore R, Kegelmeyer P 1998 Digital Mammography, chapter Current status of the Digital Database for Screening Mammography, 457–460, *Kluwer*
[14] Mari L, Carbone P and Petri D 2012 Measurement fundamentals: a pragmatic view *IEEE Trans Instr Meas* **61(8)** 2107–2115
[15] Stevens SS 1946 On the theory of scales of measurement *Science* **103(2684)** 677–680