Discussion Paper No. 120

ON THE DEFINITION OF INFORMATIONAL SIZE

by

K. Mount and S. Reiter

December 1974

ON THE DEFINITION OF INFORMATIONAL SIZE[1/]

K. Mount and S. Reiter[2/]

The purpose of this note is to explain and provide motivation for the definition of informational size given in [3]. In that paper we presented a definition of the informational size of topological spaces which defined a quasi-ordering of the class of topological spaces. (Def. 9, p. 174 [3]). In [3] we discussed certain considerations relating to the usefulness or applicability of that concept, and we also discussed certain conditions to be met by the concept (e.g., that it should agree with the ordering of Euclidean spaces by dimension); we now attempt to elucidate the informal notions that led us to this definition rather than others.

We begin with the notion of a process of observation or measurement, broadly conceived. We may think of such a process of measurement as applied to a collection of objects. It is natural to formalize this notion by means of a set S of objects and a function $f: S \rightarrow f(S)$ whose value at $s \in S$ is the measurement represented by f made on the object represented by s. Thus, if S is a collection of (names of) persons, and if the measurement process is to ask a certain question and record the answer, then $f(s)$ is the answer given by person s to the question asked; or if $g: S \rightarrow R$ represents the measurement of height, then $g(s)$ is the height of person s.

Conversely, if we are given a set S and a function f defined on S, we may regard that function as an abstract representation of an observational process on S. That is, one may conceive a procedure, not necessarily unique, which would "measure" or classify the elements of S according to the

---

1/ We have had the benefit of discussing these matters with Hugo Sonnenschein.

2/ This research was partly supported by The National Science Foundation.

equivalence relation f induces on S, (s_1 being being equivalent to s_2 if and only if $f(s_1) = f(s_2)$).

Thus, our first formal step is to identify observation processes with functions. We then consider all conceivable observation processes on a given set S, i.e., all functions on S; it is natural to interprete this as the totality of information carried by S - everything that can conceivably be observed about S.

Consider now two sets, S and T, and suppose that there are more observations that can be applied to T than can be applied to S. Since we have identified observational processes with functions, we must find a formal concept in terms of which it can satisfactorily be said that one set has "more" functions on it than another. Since the number of constant functions on a set is arbitrarily large, but not different for the two sets S and T, if there are "more" functions on T than on S, it would be natural to say that T carries more information than S, or is capable of carrying more information than S.

If S and T are finite sets, then the class of functions on each set is determined by the number of elements in S and T. Here unambiguous comparison is possible, and indeed boils down to comparing the cardinality of the two sets.

However, we want to formalize the concept of the totality of observations carried by a set so that it applies to infinite as well as to finite sets. When infinite sets are involved, because observational processes are subject to error, it is natural to consider the representation of observational processes in the context of a topological structure on the set of objects, formalizing the appropriate notion of neighboring objects, and to confine attention to observational processes which are continuous. Thus, a definition which permits us to compare the totality of (continuous) observations possible on one space with the totality of observations on another must give meaning to the idea that one topological space has "more" continuous functions on it than another. The cardinality of the set of continuous functions on a space does not provide a satisfactory concept in the case of general topological spaces. The definition we presented in [3] does give meaning to the idea that one space has "more" functions on it than another in the following way. Suppose X and Y are topological spaces and that Y is informationally larger than X in the sense of Definition 9 [3, p. 174]. According to this definition there is a continuous function $\varphi: Y \to X$ which is onto X and which has at each point of X a local inverse.

To relate this to the notion that Y has more functions on it than X, consider a measurement which could be made on X, i.e., a continuous function $f: X \to f(X) = Z$. This measurement can be "lifted" to a corresponding measurement $f^*: Y \to Z$, on Y by $f^* = f \circ \varphi$, as the following diagram shows.

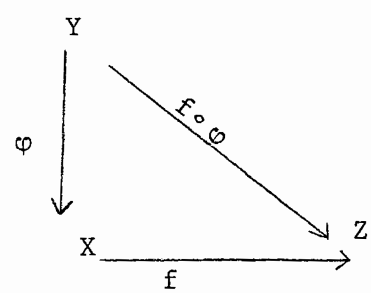

FIGURE 1.

I.e., the measurement f^* is determined at a point $y \in Y$ by first translating y into $x = \varphi(y)$ in X and then applying the measurement f to X. The result is a continuous measurement process applicable to all of Y. In this sense Y cannot have "fewer" measurements on it than X does.

[The role of the requirement that φ have local inverse has already been discussed in [3, p. 174]. We repeat here only the comment that this restriction serves to exclude "dimension increasing" continuous maps.]

Thus our definition does capture a sense in which one space can have "more" measurements on it than another. Notice that this notion is invariant under homeomorphisms of the spaces. I.e., if X has more information than Y and if X^1 is homeomorphic to X and Y^1 homeomorphic to Y, then X^1 has more information than Y^1. Thus, the quasi-ordering "has more information than" (which we may write "$\geqq$") between topological spaces, is a topological invariant.

One might be tempted to look at information about the objects being observed in terms of the observations obtained, i.e., in terms of the range space of functions. The analysis depicted in Figure 1 shows that if Y is informationally larger than X then every observation process on X can be "translated into a corresponding observation process on Y and hence that any useful purpose which would be served by making an observation on X could also be served by making the corresponding observation on Y. However notice that our definition is prior to any notion of "useful" information. For that concept a decision rule, or a preference relation, must be introduced in an appropriate way, so that certain observations would make a difference to the decisions taken or to the value of the resulting state of affairs.*/

*/ Jacob Marschak has discussed "pay-off relevant" or "useful" information in various papers, including [2]. See also [1].

One might ask about a space Y of "signals", whether it is capable of carrying the totality of information about a certain space X, representing objects. This just amounts to interpreting the inequality $Y \cdot\geqq X$ from the standpoint of Y. Similarly, to ask, what spaces X are such that measurements having values in Y can carry the totality of information about the spaces X amounts to asking which spaces X are "below" Y in the quasi-ordering of topological spaces denoted $\cdot\geqq$.

It is instructive to consider the relationship between a space X and a subspace $Y \subset X$, where Y has the relative topology. On the basis of our definition, it does not follow that X has as much information as Y. While at first sight this may appear counter to intuition, closer examination reveals that it is not an undesirable property of the definition. In terms of the informal ideas relating to observation processes sketched out above, we may notice that a subspace may be able to carry more information than the space in which it is embedded; there are measurement processes which can be carried out on the subspace, but cannot be carried out continuously on the containing space. To see this we notice first that every continuous function on the (larger) space X can be lifted to a function on the subspace Y by restriction to Y. However, there can be continuous functions on Y which cannot be extended to continuous functions on X. I.e., there are measurement processes which can be carried out continuously on a subspace of X but not on the whole of X. A simple example is provided by $Y = (0,1)$ and $X = [0,1]$, with the usual topology. The restriction of any continuous function in $[0,1]$ to $(0,1)$ is, of course, a continuous function on Y. But the function $y = 1/x$ is continuous on $(0,1)$, and has no extension to $[0,1]$. In this sense a subspace can carry more information than the space in which it is embedded. This is reflected in the ordering $\cdot\geqq$. Because the continuous image of a compact set is compact, there is no function from X onto Y, thus $X \cdot\not\geq Y$. On the other hand one may map $(0,1)$ onto $[0,1]$ continuously by a map φ carrying $(0,\frac{1}{4}]$ to the point 0, $[\frac{1}{4},\frac{3}{4}]$ to $[0,1]$ and $[\frac{3}{4},1)$ to 1. We may use as a map from $[\frac{1}{4},\frac{3}{4}]$ to $[0,1]$ a homeomorphism, from which it follows easily that φ has a local inverse.

There is reason, suggested in part by this example, to think that a local concept of informational size involving only local comparison of spaces would be useful. (See also [3] Theorem 35, p. 190.)

To express this idea we use the following definition.

<u>Definition</u> 1:

A topological space X is <u>locally informationally as large at a point $x \in X$ as a topological space Y is at a point $y \in Y$</u> if there exist open neighborhoods $N(x) \subset X$ and $M(y) \subset Y$, with $x \in N(x)$, and $y \in M(y)$, such that $N(x) \cdot\geq M(y)$.

Now, it is clearly possible (indeed otherwise there would be no point to this definition) to have two spaces such that one has locally more information than the other, but not globally. The local concept captures the idea that locally one space carries more continuous measurements than the other, but without requiring that each measurement be extendable to a continuous measurement on the whole space. The flavor of this concept may be suggested by the example of the real line and the unit circle, which are globally of different informational size, but locally the same. [See [3] <u>Example</u> p. 189].

Interpretation of the concepts of local and global informational size may shed further light on these concepts. It has been implicit in our discussion of observational processes that two aspects of them are captured in their representation by functions. The first is that ~~aspect represented by the value~~ of th function in question at a point. The second is the extension of the observation, represented by the domain of the function. In the continuous case, the idea of arbitrarily accurate identification of objects is involved in their representation as points of a topological space, as in the familiar case of objects represented by points of the real line, and the idea of arbitrarily accurate observation of them is involved in the continuity of the functions which represent observations. In this setting, the local concept of informational size may be interpreted as considering the totality of observations which are capable of being made with arbitrary accuracy on an object capable of being identified with arbitrary precision, and on all objects which closely approximate it.

The global concept of informational size involves the extension of observational processes. As we have pointed out above, the structure of relationships among the objects, as represented by the topological structure of the set of objects, can limit the extension of observational processes when we require arbitrary accuracy. Thus, the quasi-ordering of topological spaces by informational size may be interpreted as restricting attention to those observational processes which permit continuous comparison of all objects in the spaces involved.

Definition 1 can be made the basis for a concept of one space X being locally informationally as large as another space Y by requiring that for each point $y \in Y$ there exist a point $x \in X$ such that X is locally informationally as large at x as Y is at y.[*/]

---

*/ One might think alternatively of defining local comparison of X and Y by applying Definition 1 to arbitrary points $x \in X$ and $y \in Y$. With this definition it would not in general follow that if X is (globally) informationally as large as Y then X is locally informationally as large as Y, which would be anamolous.

It is clear that local comparison of spaces X and Y does not imply global comparison. However, there are conditions, which we shall not go into here, under which it is true that if X is locally informationally as large as Y, then X is informationally as large as Y.

In this brief note we have attempted to discuss some of the informal and intuitive ideas underlying our concept of informational size, ideas which we hope justify our use of the word "informational" in naming our concept. The interest and usefulness of this concept will depend on the mathematical structure it induces and on the theorems it leads to.

REFERENCES

[1] ARROW, K. J., "The Value and Demand for Information", Decision and Organization, Ed. C. B. McGuire and R. Radner, American Elsevier Publishing Co., New York, 1972, pp. 131-139.

[2] MARSCHAK, J., "Remarks on The Economics of Information", Contributions to Scientific Research in Management, Western Data Processing Center, University of California, Los Angeles, 1959, pp. 79-98.

[3] MOUNT, K., and S. REITER, "The Informational Size of Message Spaces," Journal of Economic Theory, Vol. 8, No. 2, 1974, 161-192.