THE PROBLEM OF THE FORMALIZATION OF «NEARLY ALL»

A. PHALET

1. In mathematics one has at one's disposal exact definitions of the notion «almost all» with respect to infinite sets. These mathematical definitions will form the very basis from which an attempt will be made to approach the notion «almost all» applied to finite sets. To avoid confusion the term «nearly all» shall be used instead of «almost all» when finite sets are considered.

The assertion that something is true of almost all points of a real interval (a, b) means that the set of points of which it is not true is a set of measure zero (1). When restricting oneself to the consideratin of points on the real line — this restriction does not derogate from the generality of the considered relations and properties which are not being altered when, for example, one passes on to the consideration of points on a plane — then the notion of set of zero measure can be determined in the following way:

> V has zero measure if, to each positive number ε there corresponds a denumerable collection of intervals $\{I_n\}$ which cover the set V and whose total length does not exceed ε (2), i.e.

$$V \subseteq \bigcup_{n=1}^{\infty} I_n \quad \text{and} \quad \sum_{n=1}^{\infty} |I_n| \leqq \varepsilon.$$

Let us call the notion in the above acception «almost all_1». The notion «almost all_2» can be defined in this way:

> almost all elements of a denumerable set possess a property E if only a finite number of these elements do not possess E.

The above description is a slight generalization of:

> almost all numbers of a denumerable sequence are said

to possess a property E if there is only a finite number that do not posses the property E (3).

2. We aim at a formal description of «almost all» applied to finite sets, i.e. at a formalization of «nearly all».

Let a formal description be an explicit definition in the frame of an axiomatic set theory:

$$ST \models (v)\,(Cv \rightleftarrows U), \text{ where C does not appear in U;}$$

this means that v is an instance of the defined notion C iff v can be described by means of a formula U and that every model of the axiomatic set theory ST is a model of the definition $(v)\,(Cv \rightleftarrows U)$.

The mathematical notions «almost all_1» and «almost all_2» are explicitly definable on the base of axiomatic set theory (4). Our purpose is a translaton of the given formal descriptions into definitions confined to the domain of finite sets, i.e. a transformation of «almost all» into «nearly all».

But the definitions of «almost all_1» and «almost all_2» are based upon the notion of infinite set. Consequently a simple restriction to finite collections cannot but result into a destruction of these notions. So the «formalization» of «nearly all» is beyond set theory. As will be explained this formalization requires the application of notions as neighborhood and convergence to finite sets. So, a full formalization of «nearly all» requires a «finite» or rather «relative topology» as the formal theory of the notions of neighborhood, convergence and continuity for finite sets.

In this paper we shall endeavour to translate the constituents of the notion «almost all» as it is defined in (infinite) mathematics into concepts applying to finite collections and belonging to the domain of general systems theory. The reason why the problem is tackled in this way chiefly lies in the assumption that general systems theory requires a finite or relative topology in view of a thorough explicitation of its foundations. Since such a topology is not at hand as an extensive theory one may appeal to the unanalyzed notions of general systems theory as approximations of the unknown pure concepts of relative topology. Eventually we will try to suggest how a transition from general system theory to some aspects of relative topology could be conceived.

3. First we analyze «almost all_2» and «almost all_1». Let $V_{\overline{a}}$ be an infinite denumerable set of which $\overline{a}$ represents the (infinite) cardinal number, V_n a finite subset of $V_{\overline{a}}$ with n elements.

First of all it will be investigated whether a general characterization can be given of the way in which, starting from a set V, one can get a set of almost all elements of V.

Let $\{T_i(V_{\overline{a}})\}$ be the class of the sets which contain almost all_2 elements of $V_{\overline{a}}$: $T_i(V_{\overline{a}})$ is a set of almost all_2 elements of $V_{\overline{a}}$, which is obtained by applying a transformation T_i to $V_{\overline{a}}$. We shall try to characterize the class of transformations $\{T_i\}$.

It is clear that a set of almost all_2 elements of a denumerable set is itself denumerable. So, the (infinite) cardinal a is a characteristic that is invariant for the class of transformations $\{T_i\}$. At the same time each T_i denies the invariant characteristic $\overline{a}$ to the sets V_n, where $n \geqq 1$, of negligible elements of $V_{\overline{a}}$, i.e. to the sets $V_{\overline{a}} - T_i(V_{\overline{a}})$ which are finite.

Next an interval (a, b) is taken as starting point and sets V^o of measure zero are considered.

Here $\{T_j(a, b)\}$ is the class $\{(a, b) - V^o\}$ of all those sets each of which can be said to comprise almost all_1 elements of (a, b). If $m(V_x)$ represents the Lebesgue measure of the set V_x then $m(a, b)$, $m(V^o)$ and thus $m((a, b) - V^o)$ as well, are defined [5]. We rule out the trivial case that a = b. Then it may be said that the following applies

$$m(V^o) + m((a, b) - V^o) = m(a, b) \neq 0, \text{ and since }$$
$$m(V^o) \neq 0,\ m((a,b) - V^o) = m(a, b) \neq 0.$$

In this case the invariant characteristic is the Lebesgue measure of the interval (a,b), which differs always from the measure of the neglected set V^o.

Now one can pass on to the following general formulation:

> if T(G) comprises almost all elements of G then there is a characteristic I of G, which is invariant for all the elements of the class of transformations {T} and which is no characteristic of G — T(G).

Further, if we suppose that V_n and V^o are not empty, then both classes $\{V_{\bar{a}} - V_n\}$ and $\{(a, b) - V^o\}$ are filter bases not filters, whereas $\{V_{\bar{a}} - V_n\} \cup \{V_{\bar{a}}\}$ and $\{(a, b) - V^o\} \cup \{(a, b)\}$ are filters.

As a second conclusion it can be stated that:

> if {T(G)} is the class of sets which contain almost all elements of G then {T(G)} is a filter basis and {T(G)} ∪ {G} a filter, whereas {C(T(G))} is neither.

As a third step in our analysis we aim at a formal description of the aspect of close approach which lies at the root of the specific content denoted by «almost all_1» and «almost all_2». This amounts to an examination of the relation between T(G) and G.

Instead of sets we consider characteristic functions of sets. The convergence of a sequence of characteristic functions or Boolean convergence is obtained as a special case of classical convergence [6].

> a sequence of characteristic functions $\{f^n_c\}$ converges to f_c as n tends to infinity iff ,for each x, $f^n_c(x) = f_c(x)$ for almost all_2 n.

Hence f^n_c is an element of an open set which is a neighborhood of f_c if $f^n_c = f_c$. Therefore it seems to be impossible to analyze the relation of «closeness» between $T_i(V_{\bar{a}})$ and $V_{\bar{a}}$, i.e. in the case of the notion «almost all_2», by means of classical topology.

Sets which are close to each other can be said to be «very near» to each other. Hence one could try to analyze closeness from the point of view of proximity structures. The basic term of proximity structures is the relation «near to» (7). This relation can be defined by means of the relation of topogenous order (8). A symmetrical topogenous order on a set V is defined in this way (9): A, A', B, B' being subsets of V,

TO_1. $\varnothing <_{to} \varnothing$, $V <_{to} V$;
TO_2. if $A <_{to} B$ then $A \subseteq B$;
TO_3. if $A \subseteq A' <_{to} B' \subseteq B$ then $A <_{to} B$;
TO_4. if $A <_{to} B$ and $A' <_{to} B'$ then $A \cap A' <_{to} B \cap B'$ and $A \cup A' <_{to} B \cup B'$;
TO_5. $A <_{to} B$ iff $V—B <_{to} V—A$.

A sound interpretation of $<_{to}$ is e.g. given by «is contained in», «is contained in the interior of». Less trivial interpretations will be taken into account later on.

Starting from the following definition:

A is near to B iff $A <_{to} V—B$ does not hold (10),

we proceed to the description of the relation «A is very near to B with respect to P», in short A vn_p B, in this way:

> A vn_p B iff P(A) and P(B) and, for each $A^* \subseteq A$ such that $P(A^*)$, $A^* <_{to} V—B$ does not hold.

It follows that if A is a set of almost all elements of B, then B is very near to A with respect to the (invariant) characteristic denoted by P and A is contained in B:

> B vn_p A and $A \subseteq B$.

Notice that we only consider non trivial cases, where $A \neq B$. In the case of «almost all» P denotes either «denumerability or a Lebesgue measure $> O$.

At last we have to draw our attention to a specific feature of «almost all_1».

Let $\{I_n\}$ be a cover of the set with zero measure V^o, i.e. $\sum_{n=1}^{\infty} |I_n| \leqq \varepsilon$. Moreover, suppose that, $\{I_n\}$ being the m^{th} cover, $\sum_{n=1}^{\infty} |I_n| = \overset{\varepsilon}{I}_m$; then $\lim \{\overset{\varepsilon}{I}_m\} = 0$ if $m \rightarrow \infty$. Consequently, if I represents the length of the interval (a, b) then $\lim \{I - \overset{\varepsilon}{I}_m\} = I$ as m tends to infinity.

Since I = m(a, b) one can say that the invariant characteristic, i.e. the Lebesgue measure of the interval (a, b), is the limit to which tend the measurings of a set which comprises almost all_1 elements of (a, b). So, in the case of «almost all_1» the invariant characteristic is a limiting value.

Consequently we distinguish between two senses of «almost all»:

1. the relation of closeness between T(G) and G is expressible as the relation G vn_p T(G) and T(G) $\subset$ G;
2. moreover, P may be conferred upon T(G) as the limit of a converging sequence $\{P'_n\}$.

Notice that the specification under 1 implies our first general formulation concerning G, T(G) and the invariant characteristic I.

After this we pass on to the domain of finite sets.

4. The attribution of an invariant characteristic in the cases «almost all_1» and «almost all_2» is founded on strong assuptions: the actual infinite is at the very root of their conception. In the case of «almost all_2» the main assumption is the one of the existence of infinite sets and of their properties being currently determined. As for «almost all_1» an appeal to the measure of a set was made. The definition of the Lebesgue-measure rests a.o. upon the conception of infinite collections of (disjoint) ntervals ([11]).

The problem is now whether it is possible to make assumptions which are «equally strong» as the ones mentioned above and which allow the specification of invariant characteristics and closeness in the case of finite sets.

In order to tackle this problem we assume the possibility of a reduction from infinite to finite sets. The elements of such a finite reduced set will be called states. If one could properly distinguish between an «ordinary» finite set and a finite reduced set of states our problem would, practically speaking, be solved. Since we do not have a theory of reduced sets, i.e. an axiom system for relative mathematics, at our disposal we assume that the most fundamental aspects of such a theory are part of some relatively better known domain.

Let a theory of finite reduced sets be part of a theory of general systems (GST). Then the assumption that some finite set is a reduced set cannot make GST inconsistent if it «really» is a reduced set.

Next we try to specify reduced sets exclusively by means of notions belonging to general systems theory. We propose the following description:

> a finite set is a reduced set iff the assumption of its existence does not make GST inconsistent.

It remains to explain what could be meant by the «existence» of a set in the context of a general systems theory.

Take a finite set V_f and suppose that V_f is the reduced set of the infinite set V_{inf}. Let this mean that the elements of V_f are in some way the result of a «compression» of a finite number of collections of elements of V_{inf} which is the sum of these collections. But we can only be sure of that if V_{inf} is attainable in one way or another starting from V_f. Consequently we must try to find out whether we can associate V_f with a device by means of which V_{inf} can be regained. Now V_{inf} is said to exist relative to some set theory ST if the supposition of its existence does not make the consistent theory ST inconsistent. This is of course a rather weak condition which may be strengthened at will. Then the «compression» of V_{inf}, i.e. V_f, can be said to exist if the supposition of the existence of a device of some kind — which kind actually depends to a certain extent on V_{inf} —, which has been associated with V_f, does not make the theory of devices, namely GST, inconsistent. But the determination of the kind of device to be associated with V_f must be independent of V_{inf}. Actually, the conception of finite reduced sets must be wholly independent of whatever infinite sets. So we consider the pair (V_f,S) where S determines the kind of device concerned to the same extent as is done by V_{inf}. Additional specifications concerning e.g. the choice and number of primitive elements of the device are called requirements. Let V_{req} be the set of these requirements.

Then, that there is a device of the kind determined by V_{req} for (V_f,S) means that the synthesis problem of this «finite structure» is solvable with respect to V_{req}.

The fundamental assumption relative to reduced sets or finite structures amounts to the following:

> a finite structure is said to exist in a specific way and relative to GST iff the assumption that its synthesis problem for a specific V_{req} is solvable does not nake GST inconsistent.

So, in the first place one must have the disposal of a specification within the framework of general systems theory of the synthesis problem. Synthesis of systems, as specified by G. J. Klir ([12]), is a procedure by which an interconnection of devices of prescribed types is determined for a given behaviour or a state-transition structure. Synthesis is then a relation between two structures. We are only interested in the cases where these structures are finite.

5. Let the supposition that the synthesis problem for ((V_n,S), V_{req}) is solvd be consistent with GST; then a machine structure realizing (V_n,S) should be of type T in accordance with the requirements V_{req} and the structure S. Let M_T be a device of type T and consider the ordered pair (V_n; M_T).

Moreover we suppose that if the synthesis problem relative to ((V_n,S), V_{req}) can be considered as solved without making

GST inconsistent, this is also the case for $((V_m,S'), V_{req})$, $((V_k,S''), V_{req})$, ..., where V_m, V_k, ... are non-empty proper subsets of V_n and $S' = S/V_m$, $S'' = S/V_k$, ..., i.e. S', S'', ... result from the restriction of S to respectively V_m, V_k, ...

From the foregoing assumptions it follows that the supposition that the synthesis problem is solved for $((V_m,S'), V_{req})$, $((V_k,S''), V_{req})$, ... is also consstent with GST. But then a machine realizing (V_m, S'), (V_k, S''), ... is also of type T, i.e. one can consider the ordered pairs $(V_m; M_T)$, $(V_k; M_T)$, ... From the association expressed in those pairs of the sets V_n, V_m, V_k, ... with one and the same device M_T one concludes on granting one and the same characteristic I to V_n, V_m, V_k, ...

Just as in the cases «almost all_2» and «almost all_1» the sets $V_{\overline{a}} - V_n$ and $(a,b) - V^o$ have the invariant characteristic I, respectively $\overline{a}$ and $m(a, b)$, on the ground of an assumption granting the existence of infinite sets, so the characteristic M_T is conferred to V_n, V_m, V_k, ... on the ground of a corresponding assumption for finite structures, namely the consistent supposition that the synthesis problem is solved.

Next, following our approach to a description of the notion of closeness, it can be said that V_m, V_k, ... are close to V_n relative to I, in casu M_T, if

$$V_n \; Vn_1 \; V_m, \; V_n \; Vn_1 \; V_k, \; \ldots$$

which depending on the interpretation of $<_{to}$ can be understood as meaning that neither $V_n - V_m$, nor $V_n - V_k$, ... has the characteristic I, i.e. $V_n - V_m$, $V_n - V_k$, ... cannot be supposed to have M_T as a solution of their synthesis problem starting from V_{req} and, respectively, $S/V_n - V_m$, $S/V_n - V_k$, ..., without rendering GST inconsistent.

The definition of closeness alone is not sufficient for proving that the classes $\{V_{\overline{a}} - V_n\} \cup \{V_{\overline{a}}\}$ and $\{(a, b) - V^o\} \cup \{(a, b)\}$ are filters. It requires the specification of the respective invariant characteristics to achieve this. Our analysis should have been extended to include the formal specification of invariant characteristics of finite structures. Notice that if two sets are close to a third one they are very near to each other relative to the characteristic concerned. Thus, the sets being close to a same set form an equivalence class. For the present we propose to consider this equivalence class as the class of sets containing «nearly all_2» (analogous to «almost all_2») elements of a same finite structure.

It has been said that in the case of «almost all_1» the invariant characteristic is the limit to which tend the measurings of the sets concerned. Actually, the set V° is measured and this measure tends to 0. Let a device M_T which is in fact a class of devices, be the solution of the synthesis problem of V_n, i.e. of $((V_n,S), V_{req})$. Moreover, suppose that we have succeeded in proving that the removal of some elements of V_n to get respectively V_m, V_k, ... does not rule out M_T as a solution of the synthesis problem of V_m, V_k, ... Depending on the kind of proof given this could involve an unconstructive proof of the affirmation that M_T is a solution for the synthesis problem of V_m, V_k, ... This could in turn be represented as an infinite proces of problem solving.

This proces tends to M_T as a limit. So, the measuring in the case of «almost all_1» has been replaced by problem solving in the case of «nearly all_1». That the limit, namely m(a, b), in the first case is known means, in the case of «nearly all_1» that the problem is well defined [13].

At last one may observe that further investigation has to take into account finite structures together with the rate of change of these structures and of their elements.

6. An analysis of notions such as «almost all», «nearly all» must convey, if based upon reasonably sound principles, a means to a formal approach of connected notions such as «few», «many», ... In this case a description of these notions could be outlined in the following way.

A subset V_f of V_n is said to contain a «few» elements of V_n if $V_n - V_f$ is close to V_n, i.e. «few» makes the difference between identity and closeness (in the sense defined above) of sets. By the removal of «many» elements of a set closeness is destructed: $V_n - V_m$ is not close to V_n, where V_m contains «many» elements of V_n. This amounts to the definition of «few» as «nearly not» and of «many» as «not nearly all not». As a consequence «few» is «not many», which sounds rather acceptable. Further investigations on these lines are beyond the scope of this paper.

NOTES

(1) G. TEMPLE, The Structure of Lebesgue Integration Theory, Oxford, 1971, .p 56; P.R. HALMOS, Measure Theory, New York, 1969.

(2) G. TEMPLE, o.c., p. 55.

(3) ibid. p. 37.

(4) for example, N. BOURBAKI, Théorie des Ensembles, Paris, 1960, 1963, 1957.

(5) G. TEMPLE, o.c. pp. 91, 95.

(6) ibid. p. 96.

(7) A. CSASZAR, Foundations of General Topology, Oxford, 1963, pp. X ff and 62 ff.

(8) ibid. pp. 9, 12, 62.

(9) ibid. pp. XIII, 9.

(10) ibid. p. 63.

(11) G. TEMPLE, o.c. pp. 76 ff.

(12) G.J. KLIR, An Approach to General Systems Theory, New York, 1969, p. 212.

(13) M. MINSKY, Steps towards Artificial Intelligence, in: Computers and Thought, ed. E.A. Feigenbaum and J. Feldman, New York, 1963, p. 406.