6th International Conference on Applied Informatics
Eger, Hungary, January 27–31, 2004.

Unique Identification of States of Sets

Karl Javorszky

Institute of Applied Statistics, Vienna
e-mail: javorszky@eunet.at

Abstract

There are more distinct structures on a set than sequences of its elements, if the set's cardinality is 65 +/- 32. This fact allows understanding the logical relation between combinatorical restrictions posed by a sequence on a contemporary assembly on one hand and the combinatorical restrictions posed by a contemporary assembly on a sequence on the other hand. As the information theoretical interaction within (theoretical) genetics connects information stored in a sequence (the DNA) and in a contemporary assembly (the cell engulfing the DNA), the complex combinatorical interdependence has a far-reaching relevance. In the present paper, we discuss congruent descriptions of the state of a set enumerated both commutatively and consecutively. We address specifically application possibilities in the field of encryption.

Categories and Subject Descriptors: Combinatorics, multidimensional partitions, most probable states of a structured set, linearisation of structured sets, encryption

Key Words and Phrases: encryption, concurrent utilisation of structures and symbols to transmit messages, condensation of information (memory)

1. Introduction

We look into group relations on structured sets and use their property of uniqueness for transmission purposes. In the new approach, the arguments of the logical sentence which describes the state of the set can come in any sequence within the sentence, hence the term of "commutative information storage". In opposition to the Shannon algorithm, in this approach the *sequence* of the message carrying objects is of no relevance. The collection of partitions on a set shall be called the structure of the set. About a set of a limited size, only a limited number of distinct, nonredundant sentences can be said.

2. The steps in a communication cycle

The task of one communication cycle is:

1. the identification by Sender S of that state x of the set S_S{} of his possible messages, which he wishes to transmit;
2. mapping by Sender of x to an element i of N;
3. letting Receiver R know of i;
4. mapping by Receiver of i to an element y;
5. the identification by Receiver R of that state y of the set S_R{} of his possible messages, which he understands y to mean;
6. a handshake closing this interaction

Step 3 is by its nature subject to publicity. Encryption can take place in steps 2 and 4, where the publicly transmitted message i shall be evaluated by S and R as pointers to elements within their respective arrays of potential messages.

3. The interaction between sequential and commutative descriptions of the state of the set

So far, the *sequential* property of elements of sets has been utilised for the unique identification of messages. Indeed, this very paper, or any number, is only understandable because the alphanumeric symbols are arranged in a specific order.

Group structures are known in some fields of mathematics as *sociograms*. They are the result of repeated *partitions* conducted on the set. Partition: a logical sentence which adds up summands of a natural number can be understood to describe the state of the set by telling us something about the number and size of the chunks the set is fragmented into. There is a limited number of distinct ways one can fragment a set limited in size. E.g. 5 can be in following states: {5, 4+1, 3+2, 3+1+1, 2+2+1, 2+1+1+1, 1+1+1+1+1}. The bigger the set, the more different ways are there to desintegrate it. E.g. 6 can be in following states: {6, 5+1, 4+2, 4+1+1, 3+3, 3+2+1, 3+1+1+1, 2+2+2, 2+2+1+1, 2+1+1+1+1, 1+1+1+1+1+1}. These are called *onedimensional partitions* or just *partitions* because there is no definition for other partitions but onedimensionals.

We create *combinations of partitions* on a finite set. One will note that some combinations of partitions are redundant, namely in that case that they order the set in the same fashion it has already been ordered ("before", by a different partition.)

We can distinguish different structures on a set by intuitive means. E.g. we are able to distinguish two differing structures *A* and *B,* both on a set of 6:

A: {a,ab,ab,bc,c,c}, *B:* {a,a,ab,ac,ac,bc}.

One would translate the group structures into spoken speech e.g. as follows:
A contains 3,3 (aaa,ccc), 3,2,1 (bbb,cc,a), 2,2,1,1 (abab,cc,bc,a);

B contains two times 5,1 (aaaaa,c; aaaaa,b), and once each 4,2 (aaaa,bb), 3,2,1 (ccc,aa,b).

One can still add to these observations. For purposes of distinction, it can be sufficient to point out the most improbable aspect of the state of the set. There are economies of identification effort possible in this regard.

Important: the *nature of the symbols employed* has basically no relevance in this approach to communication. E.g., the structures A and B could also be represented as follows:

A: {q,qw,qw,wr,r,r}, *B:* {g,g,gh,gf,gf,hf}.

or as well

A: {1,19,19,95,5,5}, *B:* {3,3,31,34,34,14}

and the mapping of the message to a state of the set will yield the same result. The group structure is a pointer address and the partitions are each an index on the entries into the collection. The term *sociogram* reminds us that this technique is widely used in targeted marketing or social research fields where the the database entries (the population) are subjected to segmentation and fragmentation procedures.

Sender and Receiver can keep their messages private if they agree on a specific grammar of describing their structures A and B. As each of the describing sentences is concurrently.t., there is a touch of arbitrary in the sequence of aspects one chooses to give a description of the state of the set. Each party of S and R is free to generate (agree on) its own grammar (among the limited number of possible grammars). The grammar will rule, which aspects of the structure shall be evaluated as the most important, and shall therefore be communicated next. After this sentence, there will be an implication, as the set of alternatives shall already have become restricted. The implication from the previous sentence will help to pinpoint further the element(s) by means of this sentence, but on the other hand, the sentence coming later has also a higher degree of redundancy.

The number of possible distinct combinations of partitions on a set is a direct function of the cardinality of the set. By means of a relatively small number of objects one will be able to transmit a relatively small number of distinct messages. The task is to visualise the beans of a necklace as carriers of symbols. We can transmit messages by the *structure* of the symbols' domains (the overlap) and can send the beans in a pouch, unordered, commutatively. We talk about the *cross-sectional, homotemporal* state of the set and look into the structure of the subsets.

The description of the state of the set has an arbitrary property to it. ("Which aspect shall be communicated first?") So, there is an information content in the mapping of a set of symbols (in the description) to the state of the set. The state of the set is one specific combination of partitions and the description is also a specific combination of partitions. (We have as many sentences describing what we see as there are distinguishable states we can describe.) The total information content of the structure of the set is then dependent on the cardinality (with respect to the

number of possible states of the set) and again on the cardinality (with respect to the number of possible sentences we can map to this state).

Let E(n) denote the number of partitions of n. Let n? be the number of distinct combinations of partitions on a set of n elements.

Then

$$\mathbf{n? = E(n)^{lnE(n)}}$$

We see in the logarithmic expression the decreasing information content (and increasing redundancy) of sentences describing the state of the set. There are ln E(n) difering aspects to the description of the state of the set. Whichever aspect we pick first, its sentence will be 100% useful. Aspects coming later will have more redundancy. One can transmit n? distinct messages by means of n objects carrying symbols if the objects arrive contemporal. Each of the n? states can be matched to an element of N.

In the traditional, sequence-based communication, one uses the concept of individuals. In classical set theory, each element of the set is already unique. In that concept, the sequence of the elements is unique, as each element of the set is distinguishable and its place in the sequence is fixed. Therefore, basically, one transmits in optimal case n! distinct messages by means of n objects if the message happens by means of the *sequence* of the objects. (We transmit the necklace as a necklace, and the message would be lost if the beans arrived commutatively.)

We compare the lengths of message vectors n!{} and n?{} and find:

Table 1: Numeric relation between n? and n!

n	*max(n!, n?)*
1 <= n <= 31	n! > n?;
n = 32	n! n?;
33 <= n <= 95	n! < n?;
n = 96	n! n?;
97 <= n	n! > n?.

Norming the expression n? on n! we see:

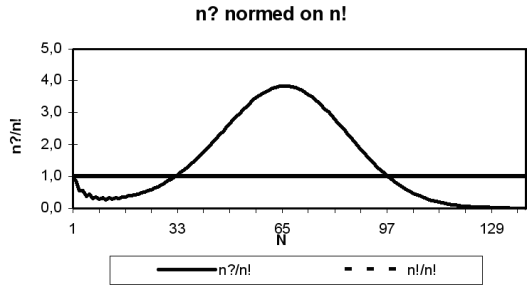

Figure 1: Number of distinct commutative states per number of distinct sequences

We interpret Table 1 as follows:

The alternatives one can read off a collection of n objects can be put to use in two basically different fashions: once as a cross-sectional, once as a longitudinal enumeration of the state of the set. In the former case, we conduct several partial enumerations, in the latter case one complete enumeration. In the former case, the disjunction among the logical arguments is the relevant aspect, in the latter case this is their distance to each other. The state of the set is mapped to N in both cases. So far, the mapping of the state of the set as a complete enumeration has been in the foreground of technical utilisation efforts. Table 1 shows that while there are more distinct states of the set as a *sequenced collection of individuals* if the set's strength is below 32 or above 97, there are more distinct states of the set as a *contemporal structured collection* if the cardinality of the set is above 32 and below 97. If the set contains 32 or 97 objects, it can be in (roughly) as many distinct states with respect to both (either) way of differentiating its elements' properties.

The graph shows that there are differing efficiency values in using a collection of objects as message carriers.

4. Discussion

The natural interdependence between number of subsets, the complexity of the overlap structures they generate among each other and the number of distinct sequences the partly individuated subsets allow on one hand and the cardinality of the set on the other hand offers very individual solutions to encryption needs of any two communication partners. There is a "natural" mapping between sequences and structures, if one uses no other symbols but as the structure itself imposes. Taking any of the natural maps as a starting-off point, Sender may veawe his sukcessive maps alternatively in an enumeration into N via n! and in an enumeration into N via n? (Step 2). The neighbourhood relations among the elements are known by S and R by their agreed-on choice of the startoff state of the set. Each subsequent communication positions the pointer in the other map of the same state of the set. The pointer having been moved in collection n? as the result of communication *i*, its position within collection n? is now the reference, relative to which the next communication *i+1* will be interpreted. Communication *i+1* moves the pointer in collection n! and its position there will then be the reference.

The neighborhood relations do need computing power to evaluate with an acceptable speed. If that requirement is met, encryption by means of using commutative states of sets in tandem with sequenced states of the set will be a piece of cake.

Next to the tandem approach there is the specialised or comment approach, too. In this case, as soon as the interpretation of a sentence in n* (n*: {n?|n!}) leads the pointer in the other language near an agreed-on edge, the language switches to that one until an agreed-on criterium is met. The common interpretation in both languages is made possible by the fact, that the more individuated subsets are generated by the structures, the more distinguishable elements are there to

sequence. There is a highly dramatic relationship between injective and surjective pictures within a structured set's structure state(s) and sequence state(s) that are congruent with each other.

References

[1] Javorszky, K.: Biocybernetics: A Mathematical Model of the Memory, Eigenverlag, Wien, 1985

[2] Javorszky, K.: Zaragoza Lectures on Granularity Algebra, Mackinger-Verlag, Salzburg, 1995

[3] Javorszky, K.: Interaction Between Sequences and Mixtures, J. of Theoretical Biology, 2000, 205, pp. 663-665

[4] Javorszky, K.: Double Sequence of Triplets,
http://www.bio.vu.nl/tmbm99/contributions.html

[5] Steidl, R. & Javorszky, K.: Messages transmission by means of counting states of sets
fis.iguw.tuwien.ac.at/resources/preprints/ karl_javorszky/MTCSS3a

[6] Javorszky, K.: Information Processing in Auto-regulated Systems, Entropy, 2003, 5, pp. 161-192

[7] Javorszky, K.: A Rational Model in Theoretical Genetics, tripleC, 2(1), 2004, pp. 20-27 triplec.uti.at/articles/tripleC2(1)_Javorszky.pdf

Postal addresses

Karl Javorszky
Landhausgasse 4/23
A-1010 Wien
Austria