# Revue Internationale de

SISGEMICITE

Vol. 4, N° **1**, 1990

afcet

Dunod

## **AFSCET**

# Revue Internationale de



volume 04, numéro 1, page 65 - 78, 1990

Reconstruction Principle of Inductive Reasoning George J. Klir

Numérisation Afscet, janvier 2016.



structability analysis is extended slightly, and a few theorems relevant to procedures for data synthesis are presented.

Although these procedures are illustrated through application to problems of decision analysis, their scope is much wider. It would therefore be worthwhile not only for its mathematical interest, but also for pragmatic reasons, to extend further the rudimentary algebra of probabilistic structure systems sketched in this paper.

#### Acknowledgements

This line of research originated in discussions with Roger Cavallo, to whom I am intellectually deeply indebted. I would also like to thank Philip Meguire for his translation into French of the abstract and for many years of discussion and correspondence that have influenced me greatly.

#### References

N. BOURBAKI, Théorie des Ensembles. Hermann Cie, Paris, 1954.

R. CAVALLO and G. KLIR, Reconstructability analysis of multi-dimensional relations: a theoretical basis for computer-aided determination of acceptable systems models, *Int. J. of General Systems*, 5, 1979, pp. 143-171.

R. CAVALLO and G. KLIR, Reconstructability analysis: evaluation of reconstruction hypotheses, *Int. J. of General Systems*, 7, 1981, pp. 7-32.

R. CAVALLO and M. PITTARELLI, The theory of probabilistic databases. *Proc.* 13th Int. Conf. on Very Large Databases (VLDB), 1987, pp. 71-81.

R. FAGIN and M. VARDI, The theory of data dependencies — a survey, in: M. ANSHEL and W. GEWIRTZ, Eds., *Mathematics of Information Processing*, American Mathematical Society, Providence, Rhode Island, 1986, pp. 19-72.

M. HIGASHI, A systems modelling methodology: probabilistic and possibilistic approaches, *Ph. D. dissertation*, SUNY-Binghamton, Binghamton, New York, 1984.

E. T. JAYNES, Prior information and ambiguity in inverse problems, *in:* D. McLAUGHLIN, Ed., *Inverse problems*, SIAM-AMS Proceedings, 14, American Mathematical Society, Providence, Rhode Island, 1984, pp. 151-166.

P. M. LEWIS, Approximating probability distributions to reduce storage requirements, *Information and Control*, 2, 1959, pp. 214-225.

M. PITTARELLI, Decision making with linear constraints on probabilities, *Proc.* 4th Workshop on Uncertainty in Artifical Intelligence, 1988, pp. 283-290.

M. PITTARELLI, Uncertainty and estimation in reconstructability analysis, *Int. J. of General Systems*, 15, 1989, pp. 1-58.

M. PITTARELLI, Reconstructability analysis: an overview, Revue Int. de Systemique, this issue, 1990.

T. SEIDENFELD, Decisions with indeterminate probabilities, *The Behavioral and Brain Sciences*, 2, 1983, pp. 259-261.

### RECONSTRUCTION PRINCIPLE OF INDUCTIVE REASONING

George J. KLIR

State University of New York <sup>1</sup>

#### Abstract

A new principle of inductive reasoning, which is based upon reconstructability analysis, is discussed. The principle differs from the straight rule, which is usually associated with inductive reasoning. Experimental studies are described by which the principle is confirmed and its domain of applicability partially delimited. The connection of the principle to the notion of pragmatic information is also mentioned.

It is assumed that the reader is familiar with the reconstruction problem of reconstructability analysis, which is overviewed in this issue by Pittarelli (1990).

#### Résumé

Nous présentons un nouveau principe de raisonnement inductif fondé sur l'analyse de la reconstructibilité. Ce principe se distingue de la règle simple, assimilée d'habitude au raisonnement inductif. Nous traitons d'études expérimentales qui confirment notre principe et qui délimitent en partie son domaine d'application. Nous signalons le rapport entre notre principe et la notion d'information pragmatique. Nous supposons que le lecteur connaît le problème de reconstruction tel qu'il se présente dans l'analyse de la reconstructibilité, résumée dans ce numéro par Pittarelli (1990).

#### 1. Introduction

During my study of the reconstruction problem (Klir, 1976, 1984, 1986), one of the two problems addressed by reconstructability analysis, I made a

1. Binghamton, New York 13901, U.S.A. Department of Systems Science, Thomas J. Watson School of Engineering, Applied Science and Technology.

Rev. intern. systémique. 0980-1472 Vol. 4/90/01/ 65 /14/\$ 3.40/@ Afcet Gauthier-Villars

surprising discovery. The discovery, which was totally accidental, occurred when I was inspecting the results of a group of simulation experiments whose purpose was to determine performance characteristics of a method for dealing with the reconstruction problem for probabilistic systems (Klir and Uyttenhove, 1977).

In each experiment, an overall system was inferred from a segment of data generated by a computer-simulated probabilistic system. This overall system was then analyzed by the reconstruction method. At each level of the refinement lattice, the method determined those reconstruction hypotheses that conformed best to the data. The conformation was imperfect in virtually all the experiments in the inspected group. In many experiments, the best conforming reconstruction hypotheses produced overall states (different in different experiments) that were not contained in the data.

When analyzing the nature of these additional states, I found, to my surprise, that almost all these states conformed to the systems by which the data were generated. That is, almost all additional states produced by the various superior reconstruction hypotheses (conforming best to data) were correct states of the system that generated the data, but they did not have a chance to enter into the rather small segments of data.

After I found similar results in other groups of simulation experiments, I started to view this phenomenon as a potential principle of inductive reasoning rather than just as an accident: the use of a reconstruction method for producing correct overall states that are not contained in the data. My initial speculations about this principle, which I named a reconstruction principle of inductive reasoning, began in the early 1980s (Klir, 1981). They were followed by rather extensive experimental studies whose purpose was to validate the principle and delimit its domain of applicability. Let me describe the nature of these studies and overview their main results.

#### 2. Experimental studies

Two major experimental studies were performed regarding the reconstruction problem: (i) an idealized study, in which it was assumed that the systems employed for generating data were structure systems (i. e., systems constructed in terms of subsystems); and (ii) a general study, in which no assumption regarding the inner structure of the systems was applied. Both studies were done for probabilistic systems; the general study was also performed for possibilistic systems. The idealized study is described in a paper by Hai and Klir (1985); the general study is described in a paper by Klir and Parviz (1986).

One purpose of these studies was to delimit conditions under which the reconstruction principle of inductive reasoning is valid. That is, the studies were intended to determine conditions under which the probabilities (or possibilities) obtained from the superior reconstruction hypotheses are better estimates of the actual probabilities (or possibilities) of the simulated system than those obtained solely from the generated data.

In each simulation experiment (of either of the two studies), three overall systems were compared: the simulated system, the system inferred from the data pertaining to the experiment, and the system reconstructed from a superior reconstruction hypothesis at some specific refinement level. Let these systems be denoted by S (simulated), D (derived from data), and R (reconstructed), respectively.

First, as part of the reconstruction problem, R was compared with D in terms of an appropriate information distance (probabilistic or possibilistic) and emerged among its competitors at the same level of refinement as a reconstructed system closest to D (best conforming to D at that refinement level). Then, R was compared with S and, similarly, D with S in terms of three criteria: an appropriate information distance, the Hamming distance between the probability or possibility distributions, and the difference between the state sets represented by the systems.

The purpose of these comparisons was to determine how the three criteria depend on four parameters: the number of variables in the systems, the number of states that the variables may assume, the size of data employed in the experiment, and the level of refinement of the reconstruction hypothesis involved. The studies were restricted (due to tremendous computational demands) to systems with five variables or less, five states per variable or less, and data with up to 2,000 observations. Thirty experiments were performed for each selected combination of values of the four parameters, and the relevant values of the information distance, Hamming distance, and state set difference obtained in each experiment were averaged over the 30 experiments.

The studies involved over 70,000 reconstruction analyses of data. Let me summarize the main results pertaining to the reconstruction principle of inductive reasoning.

The idealized study clearly demonstrated that S is always better represented by R than by D in each of the three considered criteria, provided that the data is sufficient for a reliable identification of the correct reconstruction hypothesis (the one that consists of the same subsystems as the simulated system). The reliability of identifying the correct reconstruction hypothesis was determined by calculating the percentage of relevant experiments in which

the identification was successful. For probabilistic systems, the reliability is small for small data (containing, say, less than 50 observations), but it quickly converges to 100% with increasing data size. The convergence slows down with increasing number of variables, while it speeds up with increasing number of states assumed by the variables. For possibilistic systems, the reliability is 100% (assuming S is a structure system) regardless of the data size (Klir, Parviz and Higashi, 1986).

The general study showed that the reconstruction principle of inductive reasoning is valid within a restricted domain of values of the four parameters, which depends on the risk we are willing to accept that the principle may occasionally fail. The risk is expressed in terms of the percentage of relevant experiments in which the principle actually failed. In general, the principle is applicable to data that contain no more than some specific number of observations. Let me denote this critical value of the data size by x. Clearly, x depends on the number of variables (n), the number of states per variable (s), assumed to be the same for each variable, the refinement level (r), and the acceptable percentage of potential failures (f). Although the values of x(n, s, r, f) cannot be specified precisely, they can be loosely estimated from the experimental outcomes with sufficient precision for practical applications. Considering, for example, the Hamming distance and taking n=4, s=3, f=25%, we obtain the following approximate values:

$$x (4, 3, 1, 25) = 150,$$
  
 $x (4, 3, 2, 25) = 600,$   
 $x (4, 3, 3, 25) = 550,$   
 $x (4, 3, 4, 25) = 250,$ 

x(4, 3, 5, 25) = 0 (the principle is not applicable at all).

When we apply a stronger requirement, say f = 15%, the application domain becomes smaller:

$$x$$
 (4, 3, 1, 15) = 40,  
 $x$  (4, 3, 2, 15) = 50,  
 $x$  (4, 3, 3, 15) = 100,  
 $x$  (4, 3, 4, 15) = 80,  
 $x$  (4, 3, 5, 15) = 0 (not applicable).

If we intend to apply the reconstruction principle in terms of overall states that are contained in R but not in D, the domain of applicability may be different. Let y denote the critical value of the data size in this case. This value depends on n, s, r, and the largest acceptable error (e), expressed in terms of the incorrect states in R (i.e., states that are in R, but not in S). Values of y(n, s, r, e) can be estimated from the experimental outcomes only loosely, but with sufficient precision for practical applications. For example,

when n = 4, s = 3, and e = 10%, we obtain the following approximate values: y(4, 3, 1, 10) > 2,000 (the exact value is beyond the experimental scope),

$$y (4, 3, 2, 10) = 1,500,$$
  
 $y (4, 3, 4, 10) = 30,$   
 $y (4, 3, 5, 10) = 0$  (not applicable).

Given a particular value of e, values y(n, s, r, e) represent the upper (most optimistic) bounds on the applicability of the reconstruction principle under the various values of n, s, and r. In practice, however, we are likely to be willing to use the principle only if the gain in correct states (states in R and S, but not in D) exceeds the gain in erroneous states (states in R, but not in S). This trade-off can be conveniently expressed by a simple index

$$q = \frac{|R_c| - |D|}{1 + |R| - |R_c|},\tag{1}$$

where |R| and |D| denote the total numbers of overall states contained in R and D, respectively, and  $|R_c|$  denotes the number of correct states in R. In general, q is an indicator of how much the gain (the correct novelty) exceeds the error (the incorrect novelty) when we apply the reconstruction principle. When q=1, the number of correct novel states exceeds the number of erroneous states exactly by 1 and, consequently, the reconstruction principle is applicable for situations where  $q \ge 1$ . The larger the value of q, the more powerful the principle is. As an example, Table 1 shows the values of q for n=4, s=3, and then different numbers of observations in data (N).

Table 1. Values of q given by Eq. (1) for n=4, s=3. (N: number of observations in data; r: refinement level).

| r | N = 10 | 20   | 30   | 40   | 50   | 75   | 100  | 200  | 1,000 | 2,000 |
|---|--------|------|------|------|------|------|------|------|-------|-------|
| I | 1.30   | 2.50 | 2.85 | 2.35 | 1.78 | 1.49 | 1.25 | 1.15 | 0.77  | 0.45  |
| 2 | 3.74   | 4.45 | 3.33 | 3.01 | 2.11 | 1.41 | 1.25 | 1.18 | 0.70  | 0.40  |
| 3 | 5.46   | 4.61 | 3.50 | 2.75 | 2.02 | 1.51 | 1.36 | 1.11 | 0.63  | 0.39  |
| 1 | 5.66   | 4.07 | 3.38 | 2.48 | 1.88 | 1.43 | 1.18 | 1.06 | 0.58  | 0.37  |
| 5 | 2.32   | 1.89 | 1.50 | 1.20 | 0.99 | 0.80 | 0.69 | 0.62 | 0.37  | 0.23  |

### 3. Interpretation in terms of pragmatic information

The reconstruction principle of inductive reasoning states that, under certain conditions (partially specified by the reconstruction characteristics discussed in Section 2), some justifiable novelty (something not explicitly contained in the available data) can be produced by a reconstruction method at the

expense of pure conformation to data. It seems reasonable to conceive this novelty production mechanism as follows: although the novel information is not contained in the data explicitly, it is encoded in reconstruction properties of data and can be decoded by a reconstruction method in terms of the identified superior reconstruction hypotheses at the various refinement levels.

The distance (information-based or Hamming) between R and D can be viewed in two different ways: as the loss in the degree of confirmation with regard to data and, at the same time, as the gain in the degree of novelty associated with the reconstructed system. It has often been argued that neither pure confirmation nor pure novelty contains any pragmatic information. Pure confirmation only describes the data; pure novelty basically stands for chaos. It seems that the reconstruction principle of inductive reasoning attempts to produce a compromise between the degrees of confirmation and novelty for which the pragmatic information reaches its maximum. In fact, the degree of novelty is clearly bound in this principle by the requirement that the reconstruction hypothesis with the highest degree of confirmation among a set of competing hypotheses be chosen. The actual bound depends, of course, on the level of refinement involved; in general, the higher the level of refinement, the higher the degree of novelty.

The reconstruction principle of inductive reasoning was analyzed in terms of pragmatic information by Kornwachs (1989). He argues that the degree of confirmation can be expressed by the function

where Dist denotes an appropriate normalized distance (information-based or Hamming). The degree of novelty is then expressed in terms of the same distance by the function

$$Nov(S, R, D) = Dist(S, R) - Dist(S, D),$$
  
= Dist(R, D). (3)

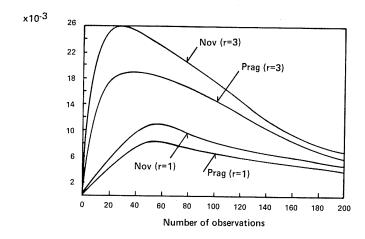
He further argues that it is reasonable to measure the amount of pragmatic information, Prag(S, R, D), obtained by the reconstruction principle by the product of the degrees of confirmation and novelty:

$$Prag(S, R, D) = Conf(S, R, D). Nov(S, R, D),$$

$$= [1 - Dist(S, R)]. Dist(R, D).$$
(4)

The knowledge of the dependence of pragmatic information on the four parameters (n, s, r), and data size), which can be obtained by simulation experiments described in Section 2, provides us with guidelines regarding the

utility of the reconstruction principle: the utility of the principle is proportional to the amount of pragmatic information it produces; its highest utility is obtained when the pragmatic information reaches its maximum.



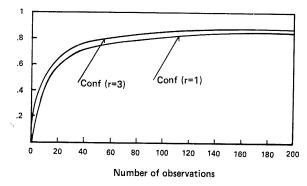


Figure 1. A typical example of functions Conf, Nov, and Prag.

As an example, Figure 1 illustrates the functions Conf, Nov, and Prag, which represent, respectively, the degree of confirmation, the degree of novelty, and the amount of pragmatic information associated with the reconstruction principle of inductive reasoning. The example is based on results reported by Klir and Parviz (1986, p. 380, Table 6) for systems with four variables (n=4), three states per variable (s=3), at the first and third refinement levels (r=1,3), and for data size range of 0-200 observations.

The plots in Figure 1 are based on 30 experiments, which were performed for each segment of data with specific numbers of observations (25, 50, 100, 150, 200 within the range shown in Figure 1). Consequently, they are only rough estimates of the functions. Nevertheless, they provide us with broad guidelines regarding the utility of the reconstruction principle, which may be described by the following fuzzy proposition: the utility of the reconstruction principle is high when the number of observations in given data is close to 50 at the first refinement level and close to 40 at the third refinement level.

We can also use the experimental results to calculate the maximum amount of pragmatic information for each data size. These maxima, together with

Table 2. Maxima of pragmatic information for n=4, s=3 (Klir and Parviz, 1986).

| Data size     | 25   | 50    | 100   | 150  | 200   | 300   | 400   | 500   | 1,000 | 2,000 |
|---------------|------|-------|-------|------|-------|-------|-------|-------|-------|-------|
| Prag          | .061 | . 044 | . 032 | .016 | . 010 | . 008 | . 003 | . 003 | 0     | 003   |
| $r_{\rm max}$ | 4    | 4     | 4     | 3,4  | 3     | 3     | 1     | 1     | 1     | 1     |

the refinement levels at which they occur, are given in Table 2. We can see that the reconstruction is best utilized when the number of observations is close to 25; on the other hand, it is not applicable or it is even counterproductive when the number of observations is close to 1,000 or greater.

#### 4. Example

A study regarding the reconstruction principle of inductive reasoning was recently performed by Hinton (1989). The study involves data consisting of a time series with 8,192 observations of two variables, each with four states. The data were collected during an ergonomic experiment; their meaning, which is not essential for our discussion, can be found elsewhere (Bayer, 1983; Klir *et al.*, 1988).

By a probabilistic mask analysis (Klir, 1985) of the data for five sampling variables, Hinton found that the mask shown in Figure 2 had the smallest predictive uncertainty. Viewing the overall probabilistic system (obtained by

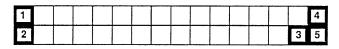


Figure 2. Mask in the example discussed in Section 4 (the integers indicate the locations of the sampling variables employed).

exhaustive sampling of the full time series by this mask) as a reference, she then studied the question of how much we can infer about this system from various subsets of the given data by using the reconstruction principle. She performed experiments with subsets of data containing 50, 100, 250, 500, and 1,000 observations. For each data size, she ran 10 experiments for different segments of the given time series and calculated average values of the individual experimental outcomes.

Using relevant Hamming distances obtained by Hinton in this study, which are not necessary to be reproduced here, we can calculate pragmatic information produced by the reconstruction principle for the various experimental instances. For each data size, maximum pragmatic information is obtained for some refinement level. These maxima are given in Table 3,

Table 3. Maxima of pragmatic information (Hinton 1989).

| Data size     | 50    | 100  | 250   | 500   | 1,000 |
|---------------|-------|------|-------|-------|-------|
| Prag          | . 007 | .019 | . 033 | . 025 | . 021 |
| $r_{\rm max}$ | 9     | 8    | 8     | 6     | 5     |

together with the refinement levels at which they occur. We can see that the reconstruction principle is best utilized in this example for data with approximately 250 observations, which is almost 25% of the 1,024 potential states ( $5^4 = 1,024$ ).

It is also interesting to examine the dependence of the number of reconstructed states on the data size and on the refinement level. The experimental outcomes obtained by Hinton are given in Table 4. The three columns under

Table 4. Reconstructed states in the study made by Hinton (1989).

| r  |    | Number of observations |   |    |     |     |     |    |     |     |     |     |       |     |     |
|----|----|------------------------|---|----|-----|-----|-----|----|-----|-----|-----|-----|-------|-----|-----|
|    |    | 50                     |   |    | 100 |     | 250 |    |     | 500 |     |     | 1,000 |     |     |
| 0  | 8  | 0                      | 0 | 15 | 0   | 0   | 36  | 0  | 0   | 60  | 0   | 0   | 95    | 0   | 0   |
| 1  | 8  | 0                      | 0 | 15 | 0   | 0   | 39  | 3  | I   | 68  | 8   | 3   | 107   | 12  | 8   |
| 2  | 8  | 0                      | 0 | 16 | 1   | 1   | 46  | 10 | 4   | 80  | 20  | 9   | 124   | 29  | 19  |
| 3  | 9  | 1                      | 0 | 18 | 3   | 1   | 54  | 18 | 8   | 89  | 29  | 17  | 132   | 37  | 32  |
| 4  | 9  | 1                      | 0 | 23 | 8   | 4   | 64  | 28 | 12  | 101 | 41  | 31  | 153   | 58  | 62  |
| 5  | 10 | 2                      | 1 | 27 | 12  | 6   | 72  | 36 | 23  | 119 | 59  | 49  | 168   | 73  | 90  |
| 5  | 12 | 4                      | 1 | 33 | 18  | 10  | 85  | 49 | 35  | 134 | 74  | 78  | 176   | 81  | 124 |
| 7  | 15 | 7                      | 3 | 41 | 26  | 17  | 101 | 65 | 61  | 164 | 104 | 141 | 195   | 100 | 214 |
| 3  | 17 | 11                     | 3 | 57 | 42  | 32  | 115 | 79 | 88  | 182 | 122 | 179 | 204   | 109 | 244 |
| )  | 21 | 13                     | 4 | 73 | 58  | 63  | 130 | 94 | 150 | 198 | 138 | 314 | 210   | 115 | 351 |
| 10 | 24 | 16                     | 8 | 80 | 65  | 125 | 134 | 98 | 260 | 199 | 139 | 543 | 212   | 117 | 649 |

each number of observations contain the following entries for each r (r=0 is used for system D):

- first column: |R<sub>c</sub>|;
- second column:  $|R_c| |D|$ ;
- third column:  $|R| |R_c|$ .

Table 5. Index q, defined by Equation (1), derived from the entries in Table 4.

|    |     | Number of observations |     |     |       |  |  |  |  |  |  |  |
|----|-----|------------------------|-----|-----|-------|--|--|--|--|--|--|--|
| r  | 50  | 100                    | 250 | 500 | 1,000 |  |  |  |  |  |  |  |
| 1  | 0   | 0                      | 1.5 | 2   | 1.3   |  |  |  |  |  |  |  |
| 2  | 0   | 0.5                    | 2   | 2   | 1.5   |  |  |  |  |  |  |  |
| 3  | 1   | 1.5                    | 2   | 1.6 | 1.1   |  |  |  |  |  |  |  |
| 4  | 1   | 1.6                    | 2.2 | 1.3 | 0.9   |  |  |  |  |  |  |  |
| 5  | 1   | 1.7                    | 1.5 | 1.2 | 0.8   |  |  |  |  |  |  |  |
| 6  | 2   | 1.6                    | 1.4 | 0.9 | 0.6   |  |  |  |  |  |  |  |
| 7  | 1.8 | 1.4                    | 1.1 | 0.7 | 0.5   |  |  |  |  |  |  |  |
| 8  | 2.8 | 1.3                    | 0.9 | 0.7 | 0.4   |  |  |  |  |  |  |  |
| 9  | 2.6 | 0.9                    | 0.6 | 0.4 | 0.3   |  |  |  |  |  |  |  |
| 10 | 1.8 | 0.5                    | 0.4 | 0.3 | 0.2   |  |  |  |  |  |  |  |

Table 5 contains values of the associated index q, defined by Equation (1), and indicates the domain of applicability of the reconstruction principle for producing novel states in this example. It should be mentioned, however, that all these calculations were made with respect to the reference system (the one inferred from the full time series), which contains only 231 states in this case. That is, states produced by the reconstruction principle that are not among the 231 states are considered incorrect. This may be questioned since it is not guaranteed that the full time series contains all states that the variables are capable of producing. In this sense, some of the « incorrect » states may be, in fact, correct predictions. Which of them to accept as feasible predictions may be decided, for example, by comparing their probabilities and accepting those whose probabilities are greater than some threshold value.

#### 5. Conclusions

Since I discovered the principle of inductive reasoning in the late 1970s, it has been an enigma to me. At first encounter, the principle has the appearance of an "information perpetuum mobile" (creating information that is not entailed in the given data) and, consequently, it is viewed with suspicion. At

the same time, however, the evidence of its success under certain conditions cannot be totally discounted. In fact, the evidence obtained by Hai and Klir (1985), Klir and Parviz (1986), and Hinton (1989) is, in my opinion, sufficiently strong to confirm the principle, even though its domain of applicability has not been adequately delimited as yet. The latter task is difficult since it requires massive computer experimentation that is both expensive and time consuming.

It is important to realize that the reconstruction principle is methodologically different from the usual conception of inductive reasoning. The latter is almost exclusively conceived in terms of the so-called "straight rule", which is well expressed by the following precept offered by Rescher (1980, p. 100):

When a certain percentage of population P have in fact been observed to have a particular trait T, then adopt *this very value* as your answer to the question: "What proportion of the entire population P have the trait T?"

The reconstruction principle violates the straight rule since it modifies, in general, the frequencies of states obtained from given data and may produce additional states that are not contained in the data at all.

How to explain the reconstruction principle? First, let us recall that the principle is based on the identification of subsets of variables of the overall system that are strongly related. These subsets of variables are expressed by the superior reconstruction hypotheses, which are obtained from the given data by a suitable reconstruction method. Assume now an investigative situation in which the data are insufficient for obtaining an adequately accurate characterization of the overall relationship among the variables, but they are sufficient for the identification of those reconstruction hypotheses that truly represent subsets of variables that, at any refinement level, entertain the strongest relationship. It is now fairly well established from the experimental studies mentioned above that situations of this sort exist for certain ranges of values of the parameters involved (number of variables, data size, etc.) and represent a window of opportunity for the reconstruction principle. Let me explain.

Since each subsystem is associated with a smaller state set than the overall state set, its relationship is, in general, better characterized by the data than the relationship of the overall system. This follows from the simple fact that the ratio between the number of observations and the number of potential states is greater for the subsystem than for the overall system. This means that the superior reconstruction hypotheses have the ability to improve our estimate of the overall relationship based upon the straight rule. Whether or not we should actually accept estimates derived from superior reconstruction hypotheses depends on the degree of our belief that the reconstruction

hypotheses in question do indeed reflect some underlying genuine reconstruction properties of the variables involved. How can the investigator of a system be helped to form rationally his or her belief in this respect? I can offer this answer: he or she can be helped by being provided with useful reconstruction characteristics obtained from suitable experiments simulated on a computer, such as those performed by Hai and Klir (1985) or Klir and Parviz (1986). Results of these experiments, when categorized by appropriate parameters, allow the investigator to compare his or her investigative situation with a comparable situation described by the characteristics and make a judgment based on this comparison. The characteristics may eventually be supplemented with appropriate guidelines of how to decide, in each particular investigative situation, whether to use the reconstruction principle or not.

An important feature of inductive reasoning based upon the reconstruction principle is that it involves aspects of both confirmation and novelty: it produces novelty and, yet, the novelty is bound by the confirmation of the reconstruction hypotheses to the given data. This seems to indicate that the investigation of the reconstruction principle in terms of pragmatic information, as initiated by Kornwachs (1989), should be very fruitful.

The problem of the justification of inductive reasoning has been a subject of great controversy among philosophers for centuries, particularly after the publication of the well-known and highly influencial analysis of the subject by David Hume (1739). Many arguments have been invented to overcome Hume's scepticism regarding the possibility of justifying inductive reasoning. When carefully scrutinized, however, each of them turns out to contain some flaws.

In virtually all attempts to justify inductive reasoning, it is taken for granted that what is to be justified is the straight rule. The reconstruction principle challenges this presupposition by establishing conditions under which it is methodologically superior to the straight rule. This extension of inductive reasoning to more than one method, each applicable under different conditions, fits well into the framework of methodological pragmatism developed by Rescher (1977). The reconstruction principle may even strengthen the pragmatic justification of inductive reasoning pursued by Rescher (1980), but it is too early to make any specific claims in this regard.

It is undeniable that major research must yet be undertaken to understand the reconstruction principle, to delimit its applicability, to develop practical guidelines of how to utilize it, and to investigate its impact on the philosophical problem of the justification of inductive reasoning. The main purpose of this paper is to stimulate the interest of researchers in various fields (mathematicians, computer and systems scientists, philosophers) to participate in this extremely challenging and potentially very important research.

#### References

- G. BAYER, Identification des menschlichen Regelverhaltens in Mensch-Machine-Systemen, *Diploma Thesis*, Institut für Regelungstechnik und Systemdynamic, University of Stuttgart, F.R.G., 1983.
- A. HAI and G. J. KLIR, An empirical investigation of reconstructability analysis, *Int. J. of Man-Machine Studies*, 22, 2, 1985, pp. 163-192.
- M. HIGASHI and G. J. KLIR, On the notion of distance representing information closeness: possibility and probability distributions, *Int. J. of General Systems*, 9, 1983, pp. 103-115.
- R. HINTON, Reconstruction characteristics of a particular probabilistic system, *Master's Thesis*, Systems Science Dept., SUNY-Binghamton, 1989.
- D. HUME, A Treatise of Human Nature, Originally published in London in 1739; a recent edition published by William Collins, Glasgow, 1962.
- G. J. KLIR (1976), Identification of generative structures in empirical data, *Int. J. of General Systems*, 3, 2, 1976, pp. 89-104.
- G. J. KLIR, On systems methodology and inductive reasoning: the issue of parts and wholes, *General Systems Yearbook*, 26, 1981, pp. 29-38.
- G. J. KLIR, Reconstructability analysis: an overview, in: Simulation and Model-Based Methodologies, T. ÖREN et al, Eds., Springer-Verlag, New York, 1984.
- G. J. KLIR, Architecture of Systems Problem Solving, Plenum Press, New York, 1985.
- G. J. KLIR, Reconstructability analysis: an offspring of Ashby's constraint theory, *Systems Research*, 3, 1986, pp. 267-271.
- G. J. KLIR and T. A. FOLGER, Fuzzy Sets, Uncertainty, and Information. Prentice Hall, Englewood Cliffs, NJ, 1988.
- G. J. KLIR, M. MARIANO, M. PITTATELLI and K. KORNWACHS, The potential of reconstructability analysis for production research, *Int. J. of Production Research*, 26, 1988, pp. 629-645.
- G. J. KLIR and B. PARVIZ, General reconstruction characteristics of probabilistic and possibilistic systems, *Int. J. of Man-Machine Systems*, 25, 1986, pp. 367-397.
- G. J. KLIR, B. PARVIZ and M. HIGASHI, Relationship between true and estimated possibilistic systems and their reconstructions, *Int. J. of General Systems*, 12, 4, 1986, pp. 319-331.
- G. J. KLIR and H. J. J. UYTTENHOVE, On the problem of computer-aided structure identification: some experimental observations and resulting guidelines, *Int. J. of Man-Machine Studies*, 9, 5, 1977, pp. 593-628.
- K. KORNWACHS,, Reconstructability analysis and its re-interpretation in terms of pragmatic information, in: *Advances in Computer Aided Systems Theory-EURO-CAST'*89, R. MORENO-DIAZ and F. PICHLER Eds., Springer, Heidelberg and New York.

REVUE INTERNATIONALE DE SYSTEMIQUE Vol. 4, N $^{\circ}$  I, 1990, pp. 79 à 84

M. PITTARELLI, Reconstructability analysis: an overview, Revue Internationale de Systémique, this issue, 1990.

N. RESCHER, Methodological Pragmatism: A Systems-Theoretic Approach to the Theory of Knowledge, New York University Press, New York, 1977.

N. RESCHER, Induction, Basil Blackwell, Oxford, 1980.

### QUELQUES REMARQUES SUR LA THÉORIE GÉNÉRALE DES SYSTÈMES ET SES APPLICATIONS

John ROSE

World Organisation of Systems and Cybernetics 1

La conception la plus répandue concernant la « science appliquée » est celle de la maîtrise de l'environnement en vue d'une fin déterminée : faire en sorte que des choses se produisent ou empêcher que des choses se produisent. Avant l'avénement de la science, la magie (pensait-on) aidait à parvenir aux fins en question. Associé à l'espoir de maîtriser les événements se trouve l'espoir d'au moins les prévoir, puisque la prévision permet à l'homme d'adapter ses plans et ses actions.

La science a rendu possibles à la fois la prévision sûre et la maîtrise efficace des événements naturels (ou du moins de certains d'entre eux) et cette connaissance a fondé le lien intime entre la prévision et la maîtrise grâce au paradigme fondamental de l'assertion scientifique: « si ..., alors ... ». Des propositions de ce genre, que l'on peut justifier par des prévisions vérifiables et/ou par des actions efficaces, sont généralement considérées comme constituant la science elle-même. Cette conception est sous-jacente aussi bien à l'attitude positive qu'à l'attitude négative envers la science. L'attitude positive, engendrée par la réalisation de rêves millénaires, n'exige aucune explication. L'attitude négative est engendrée par l'échec de prévisions ou d'actions (échec largement considéré comme caractéristique de toutes les tentatives de création d'une science sociale), par l'usage de la maîtrise de forces naturelles à des fins de destruction (par exemple dans le cas de la « mégatechnologie » militaire), par l'existence de sous-produits indésirables de la domination exercée sur la nature (dégradation de l'environnement), ou enfin par des usages imaginés de la prévision et de la maîtrise dans le but de dominer les masses (lavage de cerveaux, ingénierie génétique).

1. 5, Margate Road, St Annes-on-Sea, Lancashire, FY8 3EG, Grande-Bretagne.

Rev. intern. systémique. 0980-1472 Vol. 4/90/01/ 79 /06/\$ 2.60/© Afcet Gauthier-Villars