N90-29043

MULTIRESOLUTIONAL MODELS OF UNCERTAINTY GENERATION AND REDUCTION

A. Meystel
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, PA 19104

Abstract
Kolmogorov's axiomatic principles of the probability theory, are reconsidered in this paper in the scope of their applicability to the processes of knowledge acquisition and interpretation.The model of uncertainty generation is modified in order to reflect the reality of engineering problems, particularly in the area of intelligent control.This model implies algorithms of learning which are organized in three groups which reflect the degree of conceptualization of the knowledge the system is dealing with.It is essential that these algorithms are motivated by and consistent with the multiresolutional model of knowledge representation which is reflected in the structure of models and the algorithms of learning.

Key Words: Abstraction, Generalization, Error, Interpretation, Knowledge, Model, Multiresolutional, Redundancy, Representation, Tier, Uncertainty.

1. Introduction

Multiresolutional system of Knowledge Representation (MKR) and processing (MRKP) is based upon five postulates formulated in [1]: (P1) - descend from the verbal descriptions, (P2) - existence of the external global thesaurus with interpretations, (P3) - dependence on the context, (P4) - metrizability, (P5) - holism[1]. All of these postulates establish representation as a body which must be uncertain. Indeed, the set of verbal descriptions which is the source of representation cannot be complete, and all of these descriptions cannot be adequate (uncertainty of incompleteness and of inadequacy), interpretations from the global thesaurus can be utilized only if they are curtailed (uncertainty of abridgment), context allows for subjective processes encoding and decoding (uncertainty of subjectivity), metrizability is possible only within a certain scope of consideration

[1] A single tesselatum cannot be used for representation, only a complete set of all tesselata can represent the system, the sets of mechanisms of generalization operating among the tessellata is also a part of representation; thus redundancy of representation cannot (and should not) be avoided.

(uncertainty of scoping), finally, the mechanisms of generalization and instantiation carry with themselves all four sources of uncertainty mentioned above, and yet we have to use them (uncertainty of inference [2]).

Uncertainty calls for evaluation which is required for decision making. Indeed, after the alternatives of the future decision are constructed (whether in the problems of design, or in the planning/control problems) these alternatives are to be compared. Consistent comparison can be done only if the judgment is developed about the uncertainty of the evaluation of our alternatives. Each alternative together with the evaluation of its merits and shortcomings has a definite probability of occurrence. The set of alternatives is obtained presumably by combinatorial methods discussed in ATG area [3,4]. The combinations will be compared based on a set {merit, shortcomings, probability]. Thus, the body of probability theory should be evaluated in order to answer the question: can we use its recommendations in the process of uncertainty evaluation?

This makes the 6 famous Kolmogorov's axioms [2] a mechanism that can be used for making our judgment on the utilization of the theory of probability per ce. His axioms are stated for the set E of elementary events which are called elementary events ($E=\{\psi, \eta, \xi, \ldots\}$, F is the set of subsets of E, and elements of F are called random events. These are the axioms formulated by Kolmogorov for the system consisting of E and F.

Axiom 1. F is a field of sets [3].

Axiom 2. F contains the set E.

Axiom 3. To each set $A \supset \mathsf{F}$ is assigned a non-negative real number P(A). This number P(A) is called the probability of the event A[4].

Axiom 4. P(E) equals 1.

Axiom 5. If A and B have no element in common then $P(A+B)=P(A)+P(B)$.

The set of couples $\{\mathsf{F}, P(\mathsf{F})\}$ is called a *field of probability* where $P(\mathsf{F})$ are the probability values satisfying Axioms 1-5.

Axiom 6. For a decreasing sequence of events $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ of for which the

[2] It is presumed that inference is built upon parallel or sequential mechanisms of generalization and instantiation, (easy to verify, all known rules of inference and logical resolution are based upon determining properties of belonging to a class, or forming a class).

[3] Field is understood as a system which includes all sums, differences, and products of all elements as well as all subsets of it. So F is understood as a mechanism of generating combinations.

[4] One can also use the words: "possibility", "preferability", and so on. The idea of relative frequency is never raised in the set of axioms. This means that the axioms may fit into the structure of fuzzy set theory, Dempster-Shaffer theory, and so on.

product of all sets $\prod A_n=0$ the following equation holds: $\lim P(A_n)=0$, if $n \to \infty$.

We would like to question the validity of the Axioms of 5 and 6 for the case of MKR. Indeed, the phenomenon of having no element in common is not a simple thing especially taking in account the fact that even within a single tessellatum all objects under consideration can be often considered as built of the same primitives (components). On the other hand, condition $\prod A_n=0$ means that the events (sets) under consideration are incompatible. However, the infinite inclusion $A_1 \supset A_2 \supset \ldots \supset A_n \supset \ldots$ does not require necessarily that $\lim P(A_n)=0$, if $n \to \infty$. Everything depends on interpretation of inclusion. This becomes especially important when the process of consecutive generalization is considered.

As it was mentioned in [1], the core of MRKP operations does not differ from the process of the **automated theory generation (ATG)** [3,4]. Since the mechanisms of generalization are involved, then any **process of representation is based upon theory generation**. Like in ATG, the subsystem of representation is supposed to invent and utilize an algorithm of transforming a tessellatum built at a definite resolution into tessellata of lower resolutions. This can be considered a process of synthesizing a consistent system of tessellata constructed at different resolutions and transformable one into another. This synthesis can be performed in a different way depending on initial problem specifications, and entail different results. So, MKR is a source of uncertainty which cannot be considered a fault or a failure: this is an intrinsic property of the system which should be properly understood rather than to uncompromisingly fought with.

2. General Mechanism of Knowledge Processing (GMKP).

A structure of GMKP is demonstrated in Figure 1. It operates as follows.

1. A subset of an object is considered to be of interest. It is presumed that this sub-object (SO) is a part of an object, which in turn, is a part of a particular Domain, which finally, is a part of the World[5]. Information concerned with SO (ISO) is obtained through the set of available sensors which can include all practical variety of them starting with the transducers for delivering actual physical information transformed into a form convenient for the particular system configuration, and ending with the terminals for computer reading necessary documents.

[5] It is important to accept the existence of the World as a part of the problem even is the problem is specified within an extremely narrow domain with a small subset of an insignificant lonely object. The World affects the problem in a powerful way almost in all known cases: via thesaurus, and the process of interpretation no operation of MKR can be performed without taking in account the links with the World.

The Sensor Information Carrier (SIC) delivers ISO to the system for MRKP in a form that contains information about the *code* carried by this particular SIC, and about the *modality* of this particular sensor. The code contains the information of the label and the value, this information should be decoded, and the process of inference is performed, after which all information is structured and stored which (after this) makes it *knowledge* [6].

The left part of the Figure 1 can operate only if the right part exists. As soon as the modality[7] of sensor information is becoming known, o particular Domain of the World Knowledge is being evoked, and the mechanism of interpretation is being prepared taking in account the context, and listing the available rules that can be utilized by the system for dealing with the decoded information.

After the Storage of Knowledge received the interpreted information, the mechanisms of learning are getting involved. The whole body of the stored knowledge is reconsidered in the view of correctness of the classification results after the new information has arrived. As a result of the learning process, new rules for interpretation can be obtained which in fact can affect the process of interpretation and inference and change the prior (recent) results.

SO generates all sources of uncertainty: error of measurement (E), uncertainty of incompleteness (I), and uncertainty of redundancy (R). New EIR-uncertainties are generated within the code as a result of coding and communication; within the interpretation as a result of the EIR-interpretation properties, and within the storage as a result of EIR-properties of classification and other tools of information organization. On the other hand, all subsystems of the right part of Figure 1 contain the same deficiencies.

All these factors should be taken in account when the degrees of belief are being determined. Usually they are generated within the loop of "learning - interpretation - storage". This is why we are especially concerned with the EIR-properties of the external bodies of knowledge which are used for interpretation. One of these properties is the frequency of updating. The case presented in Section 3 should illustrate how these properties are being generated.

[6] Knowledge is defined as internally structured information considered to be a part of some external organization, and allowing for interpretation in some particular context.

[7] Modality of sensor is understood as a subset of the physical phenomena this sensor can sense and submit to the system (like vision, hearing, touch (i.e. surface properties are being sensed), temperature,and many others).

3. A Case Study "Knowledge of a Particular Actuator".

We will discuss knowledge of a particular type of machine actuator: induction motor (the results may be partially used for the similar analysis of synchronous, and DC brushless motors). The model in a form of system of differential equations is very complicated for all these types of actuators, also it is inconvenient in practice of utilization, and can generate many errors because of errors in input information, and because of many factors one neglects in order to use differential equations for modeling the induction motor).

In Figure 2,a a typical "speed-torque" curve is shown which in decades was used to describe the operation of induction motor. Analytically it can be represented in a form

$$T(s) = \frac{2T_{max}}{\dfrac{s}{s_{max}} + \dfrac{s_{max}}{s}}$$

where $s=(\omega_0-\omega)/\omega_0$ is a so called "slip" (difference between the speed of the rotating field and the speed of the rotor), ω-speed, T-torque, T_{max}-maximum torque, s_{max}- "slip" corresponding to the maximum torque. This formula (first derived by Kloss) was successfully used in decades. About half a century ago, some reservations were voiced. The Kloss formula was good when its correctness was verified by measurement performed by the Prony method using a very imprecise and often messy method of measurements based upon lever with friction balanced against the torque developed on the shaft. More accurate measurements performed by 2 and 3-machine aggregates led to the experimental data which looked like a curve shown in Figure 2,b.

It was clear that there are many factors creating the phenomenon of these "distortions), and the researchers started working on this "enigmatic"[8] behavior. In forties it became clear that instead of the Kloss formula one should use an expression which looks like

$$T(s)=\sum T_i(s)$$

which contains many Kloss formulas for a variety of the following factors:

1) Fields of "teeth" harmonics (those substantial in magnitude),
2) Imaginary "skin-cylinder" rotor which appears because of the final

[8] Enigmatic-usually means: not coinciding with the model I (he, scientific community) thought of.

surface machining of the rotor,

3) Components of the torque which are generated in the zones of bad insulation among the laminations of the rotor "iron",
4) Harmonics of the stator spatial field due to the nonsymmetrical distribution of the stator winding along the inner surface of the stator window,
5) Saturation of the machine,
6) Nonsymmetry of the stator voltage, and so on

....

Using all of them in all cases would be totally senseless. Using some of them in analytical form would be a matter of choice for a particular designer. Manufacturers started giving instead of a curve T(s) a fuzzy zone around the imaginary average T(s).

In the fifties the topic with T(s) has been exhausted, but one still could not compute a system with induction motor. Now it was clear that something else generates error: probably the dynamic processes. Thus the focus of attention of the researchers in this area shifted toward the time diagrams of speed and torque. Full twelve dynamic equations of the three-phase induction motor (highly nonlinear and coupled) were hard to use, and not always easy to believe. In Figure 3,a two curves demonstrate: 1) T(t) if the Kloss formula is used, 2) T(t) often seen in practice (the phase portrait is shown in Figure 3,b). In the sixties it became clear that due to the exponential components of the current in the winding at the moment of connection to the line, a "swinging" component of the field generated the oscillations with the frequency of the voltage[9]. Oscillations in the end of the process were on the natural frequency of the motor (if modelled as a second order system). It the systems with SCR switches and/or controllers the first component could be controlled and even eliminated. (The whole picture is even more complicated, but the major factors here are presented properly). In Figure 3,c a set of voltage and harmonics is shown dependent on the components of a real speed-torque characteristics.

In Figure 4 the whole multiplicity of factors to be taken in account is collected in a hierarchy. The lower is the level of consideration, the more one can find items (components of the torque) which can be neglected under proper circumstances, which are not as significant as the other components that are retained when the information is generalized for submission to the upper (low resolution) levels. However, it is clear that no judgment of error can be done unless the system is considered as a hierarchy of generalizations and its model can be discussed tessellatum by tessellatum together with their inclusion rules [1].

The following observation can be formulated for each two adjacent resolution levels (a

[9] As a matter of fact, this component was to blame for 70% of all shaft breakages known in industrial practice.

tier) which is being confirmed by the variety of other technological examples. Learning process consists of consecutive refinement of the sets of knowledge containing error [10] (for the low resolution level of all tiers) with transforming it into goal oriented modelled knowledge (for high resolution levels of all tiers). The following questions must be addressed before the qualification of data is done as containing some error and evaluating this error.

1. What should have been considered an error at each stage of our development of the model of T(s) a) a model error, b) an error of measurements ?

2. Which of the models should have been considered a "true model" for dealing with the problem of error evaluation and qualification?

4. Model of Error Generation and Reduction

We will address these questions in a form of recommending a general approach for dealing with processes of error propagation in the system, and for recommending measures of its reduction. Let us first consider the updated Figure 1 which can be corrected based upon the principle of learning formulated above (see Figure 5). We saw that no judgment of error can be made before the system is organized as a hierarchy of generalizations (abstractions, multiresolutional hierarchy, etc). Thus, the hierarchy of resolution conscious information should be sought for from the SO. Then, the Code which arrives should be considered a generalized code which allows for nested hierarchical treatment (recursive interpretation, and/or consecutive refinement). Thus, in Figure 5 a loop is shown from Generalized Interpretation (GI) back to Generalized Code.

Now, the storage is becoming a multiresolutional system, and the whole right side of the structure is being adjusted with methodology of [1]: the source of knowledge is being treated as a multiresolutional structure, rules constitute a hierarchy of classes and a hierarchy of rules within the class, finally, the processes of learning are done consecutively with gradual involvement of each consecutive tessellatum. The system of learning is shown in Figure 6.

Then the following conceptual structure is required to support the MRKP system in the view of dealing with processes of error generation and its reduction (Figure 7). The whole processing is considered as a multiresolutional system of consecutive encoding/decoding procedures. In a number of cases a hierarchy of sensors can be expected that makes the encoding subsystem working with a multiplicity of inputs to all levels.

The process of consecutive refinement is illustrated in the structure of search shown in

[10] Error is understood here as a deviation from experimental data.

Figure 8 for $\sqrt{2}$. In this case the two conditions are to be satisfied

1) $[(\text{upper level})+(.1)(\text{lower level})]^2<2$;

2)

$$(\text{lower level}) < \frac{2 - (\text{upper level})^2}{(.2)\ (\text{upper level})}$$

Each upper level is obtained by averaging the lower level. This generalization rule: averaging is expected to be domination for the highest levels of resolution. It allows for dealing only with interval type of the error with uniform distribution of error within the interval.

The following conclusions can be made:

1. The characteristics of error will depend completely on the procedure of generalization accepted within the particular paradigm. Averaging seems to be the most appropriate procedure of generalization for the higher levels of resolution.
2. Instead of dealing with the second order statistics one can deal with a resolutional hierarchy of first order statistics, each of them for the interval error with a uniform distribution within the interval.
3. If the nature of the error allows for possible models of errors with the infinite interval, it can be substituted by the interval error with the same entropy of the error.
4. Errors are to be dealt with using algorithms of Multiresolutional Nested Consecutive Refinement.
5. Information improvement (learning) procedure can be arranged which allows to predict the level of uncertainty, and to postpone the decision making until the desirable level of uncertainty is achieved.

References

1.A.Meystel,"Techniques and Potential Capabilities of Multiresolutional Information (Knowledge) Processing" (in these Proceedings)

2. A. Kolmogorov, Foundations of the Theory of Probability, (first published in 1933 in Germany, in 1936 in Russia, in 1950 in USA), Chelsea Publishing Company, NY 1956

3. S. Amarel, In Principles of Self-Organization, Trans. of the University of Illinois Symposium on Self-Organization, June 8-9, 1961, Eds. H.Von Foerster, G. Zopf, Jr., Pergamon Press, Oxford, 1962

4. R. Davis, D. Lenat, Knowledge-Based Systems in Artificial Intelligence, McGraw Hill, NY 1982

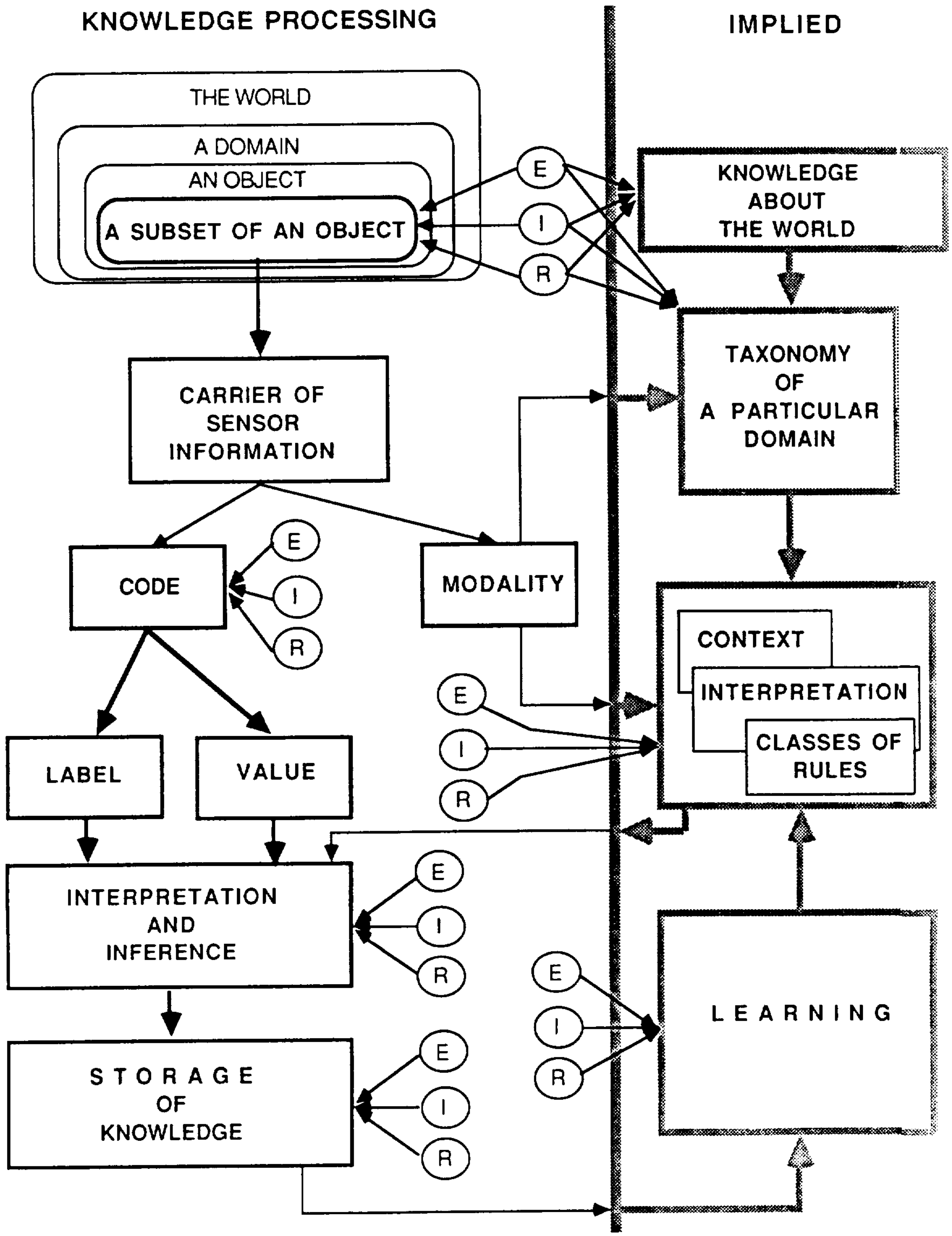

Figure 1.Conceptual Knowledge Processing diagram, and the implied bulk of the required external knowledge support.

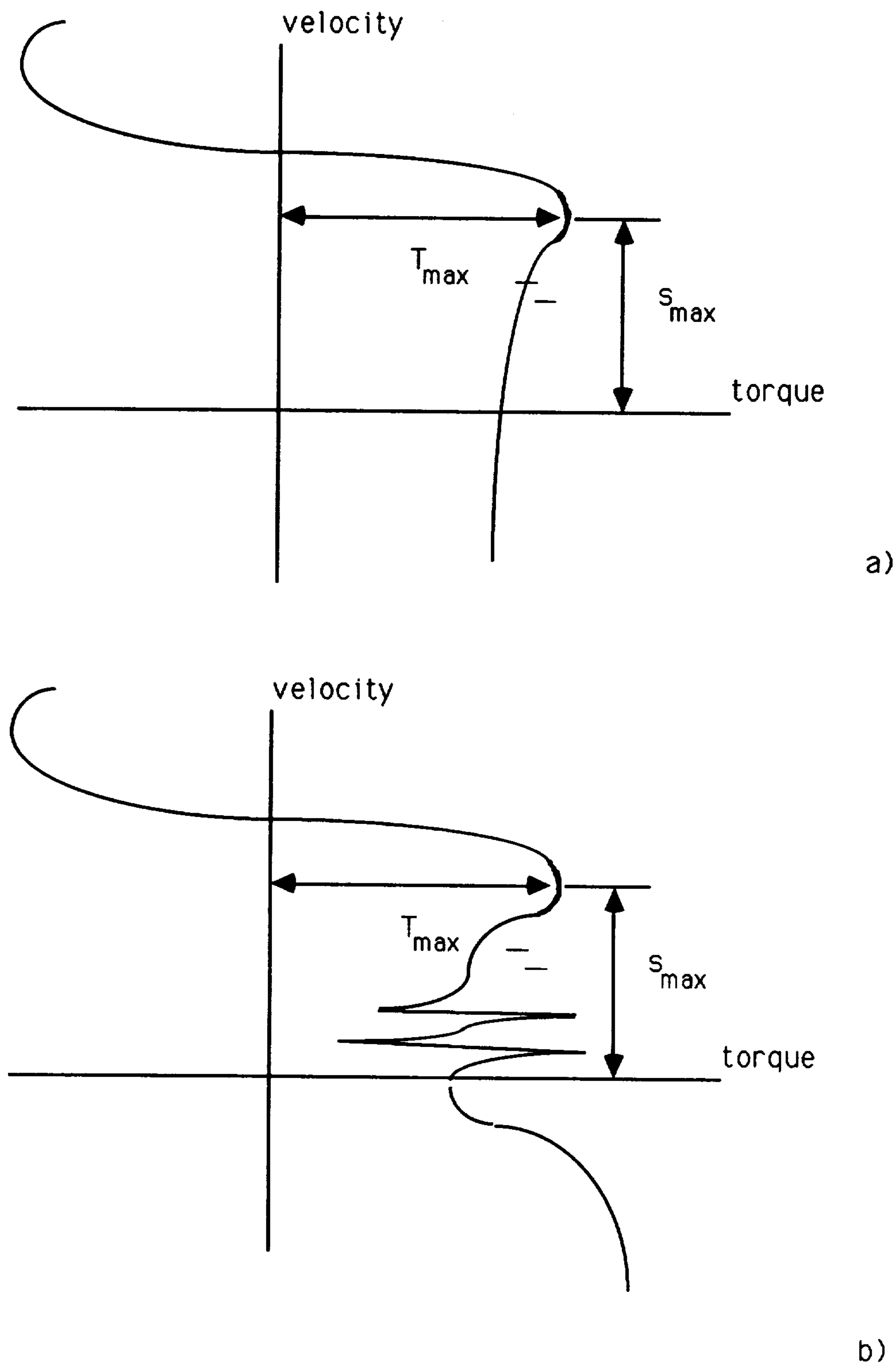

Figure 2. Torque-speed characteristics of an induction motor

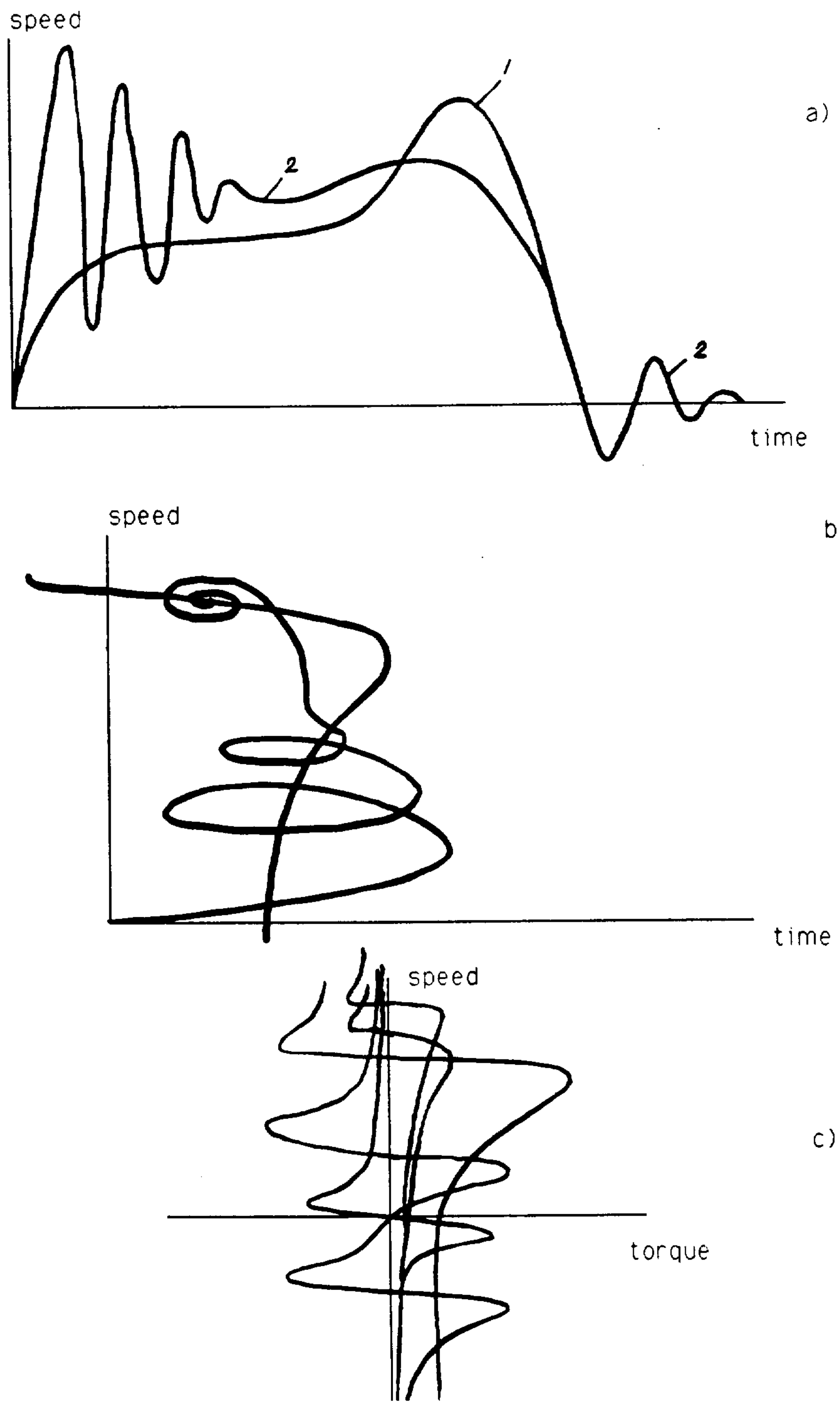

Figure 3. Dynamic Characteristics of the Induction Motor

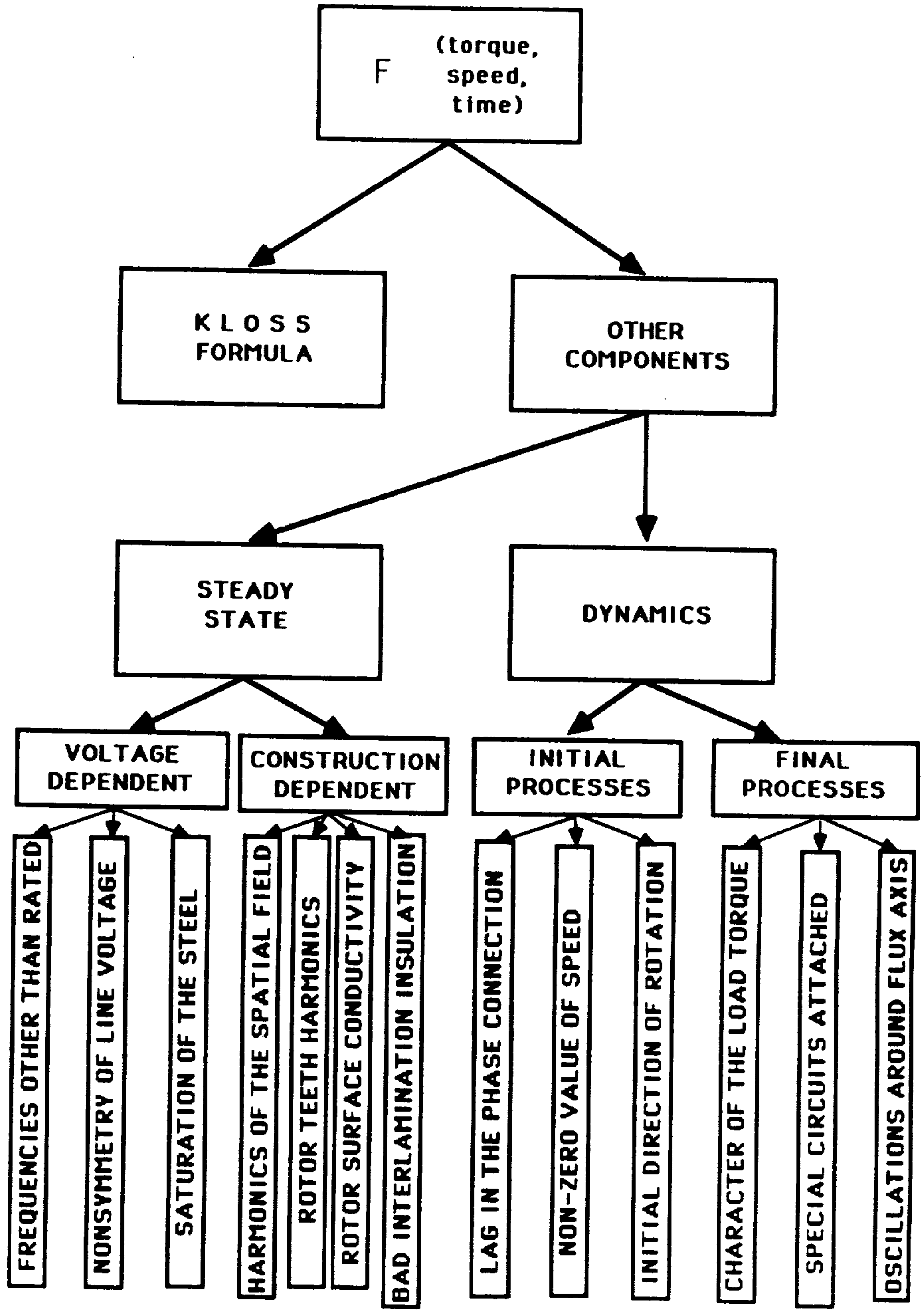

Figure 4. Structure of knowledge refinement for the case with induction motor

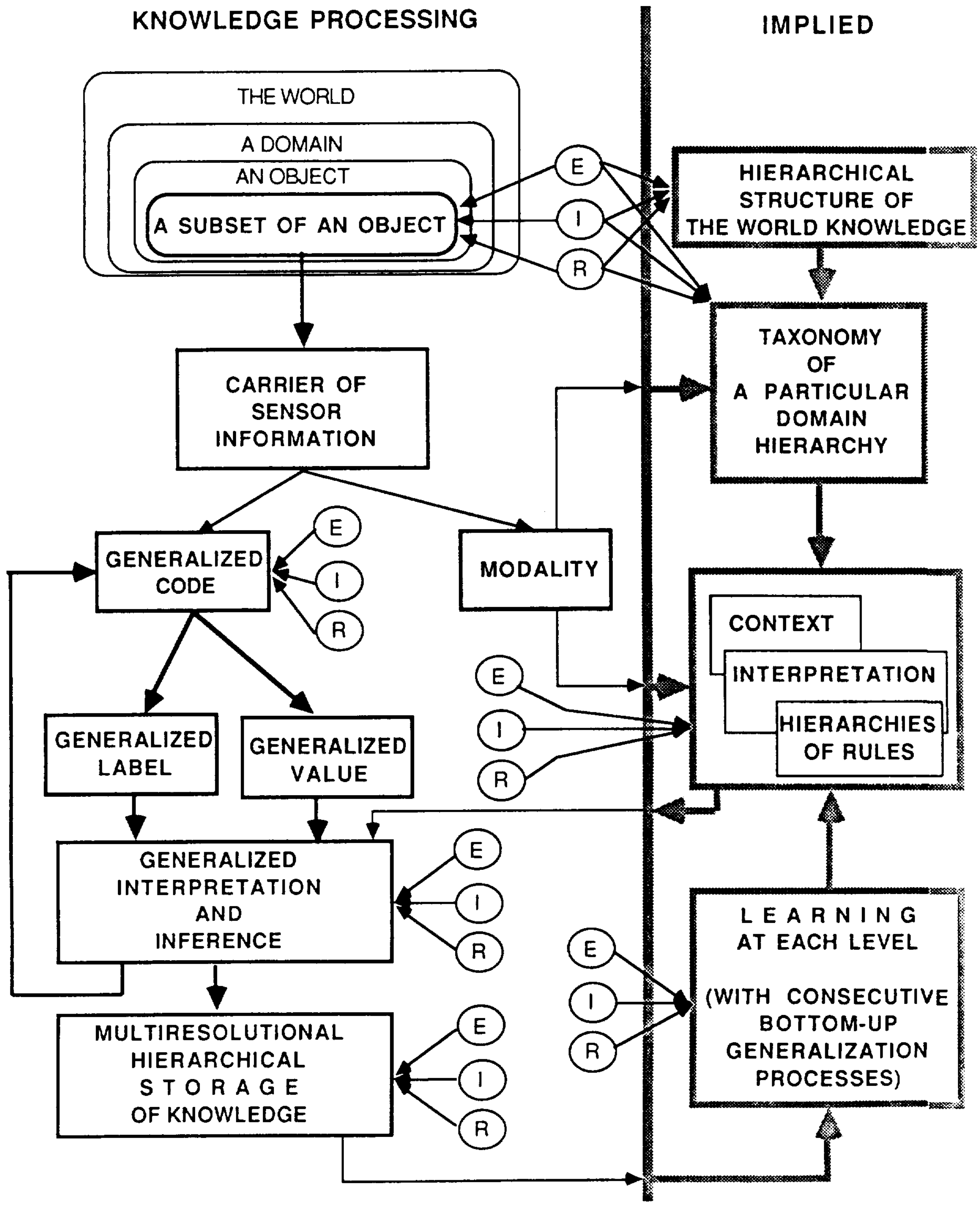

Figure 5. Multiresolutional Knowledge Processing (MRKP) diagram, and the implied bulk of the required external knowledge support.

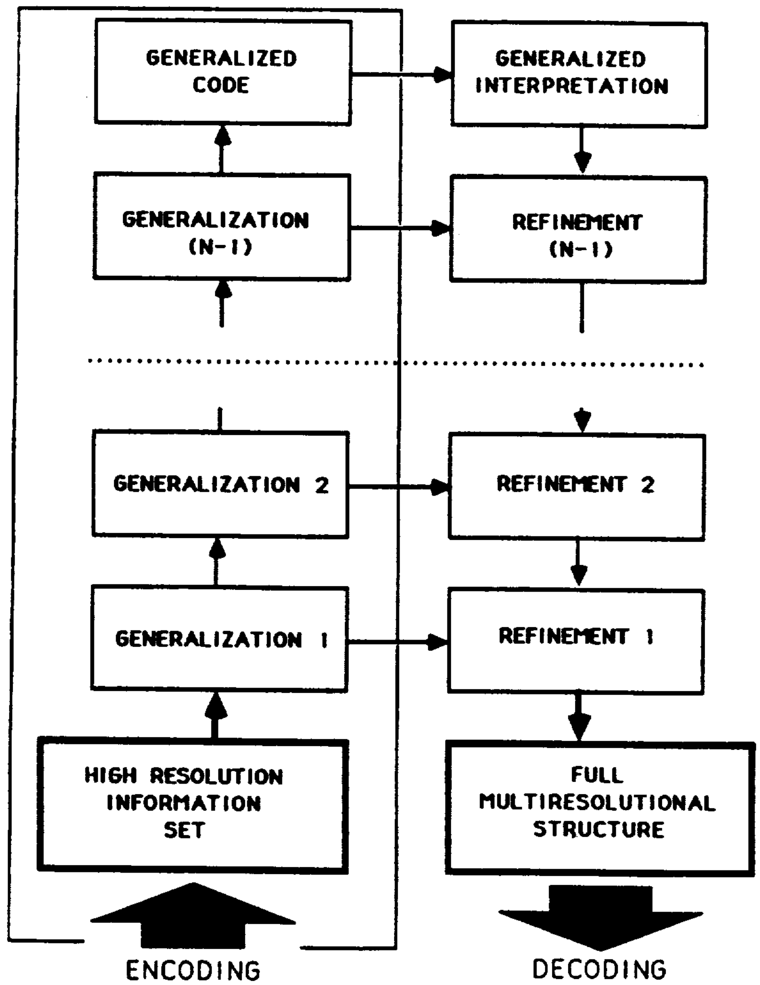

Figure 6. Mutually supportive processes of "encoding-decoding"

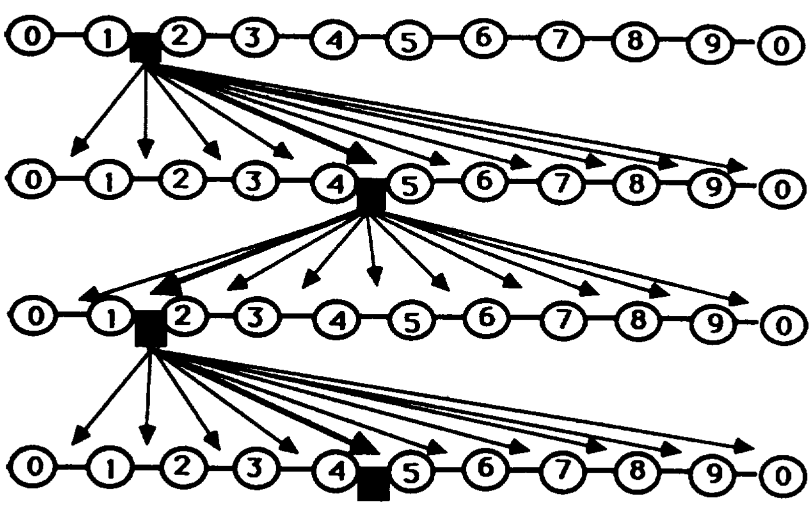

Figure 7. MRKP for solving $\sqrt{2}$ problem