Towards a Support Framework for the Generation and Exploration of Conceptual Design Solutions

Amaresh Chakrabarti
EDC, Department of Engineering, University of Cambridge, Trumpington Street
Cambridge CB2 1PZ, UK [ac123@eng.cam.ac.uk]

SUMMARY

In this article, existing literature reporting the evolution of micro-sensor concepts is reviewed to propose a framework for supporting such development processes. The framework is based on a design-evaluate-modify cycle where designing is done at three abstraction levels: device functionality, causal networks of physical effects as solution principles which can satisfy the functionality, and networks of physical structures as physical concepts that embody the solution principles. Based on the available expertise between the EDC and RACE, a preliminary research plan for the development of the framework is proposed, and a research method is presented. Following this, a brief overview of the existing approaches to the generation of design concepts in the EDC and RACE is presented. Finally, the applicability of these approaches to micro-sensor designs is discussed.*

1 A CONCEPTUAL DESIGN EXAMPLE: SOME OBSERVATIONS

Let us take a snapshot of the development process of one of the concepts for a micro acoustic sensor [taken from Sessler, 1994]. A concept based on the condenser principle was thought of, where the sensor consisted of a backelectrode chip and a membrane chip glued together with a 2 μm air gap separating the backplate and membrane. The idea was that membrane deflections, caused by the pressure of sound waves, would change the air gap, thereby causing the capacitance to change, which could be sensed in the electric circuit. However, apart from causing a change in capacitance across the air gap, the membrane deflections also cause streaming of air in and out of the air gap, thereby introducing damping. This in turn causes a loss of sensitivity of the sensor at highter frequencies.

The drawback of air gap damping, which is typical of condenser microphnes, can be eliminated in various ways. One approach utilises the fact that this damping is inversely proportional to the third power of the air-gap thickness, and thus an increase in thickness will result in a much reduced damping. Another approach to reduced air-gap damping is to perforate the backplate with a large number of holes connecting the air gap with a large air cavity.

It is to be noted that all the above modifications are obtained by tinkering with essentially the same central solution principle: sensing a change in capacitance caused by a change in air gap due to a deflection of one of its plates caused by sound-pressure. Alternatively, one could dream up other ways of sensing acoustic pressure. For instance, these sensors can be, and are, developed on the principle that sound waves cause a phase or intensity modulation of light waves in an optical fibre. The concepts based on this new principle, as in the case of the condenser concepts, would have their specific problems. For example, because of the strong temperature dependence of light propagation in a fibre, phase modulating transducers tend to instability and require compensating measures.

Let us counter this example with the problem of high-resolution surface probing. One of the principle solutions to this problem is to move the tip of a cantilever beam, which is in contact with the surface being probed, across the surface; the deflection of the beam can now be measured using a number of principles, including those mentioned above [Brugger et al., 1993].

A number of issues could be observed from the above descriptions, which are listed below.

• For a given design problem, a conceptual solution principle can be described as a causal network of physical effects. For instance, in the condenser principle, a change of pressure causes a change of force, the change of force causes a deflection, this deflection causes change in an air gap, the change in the air gap causes a change of capacitance, etc. Together they form a causal chain of a set of physical effects such as stiffness etc to provide the overall input-outputs of the problem.

* Several ideas presented in this article were stimulated by discussions with Prof. Taura and Prof. Kiriyama during the author's visit to RACE in August 1994, and is thankfully acknowledged.

• Usually there are more than one central solution principle which can be used to solve a given design problem. For instance, acoustic sensors can be based, among others, on the condenser principle or on the optical phase or intensity modulation principles. This means that if a set of the physical effects were available in an appropriate form, these can be woven together into alternative causal networks as alternative solution principles to given design problems.

• Usually there are several ways in which the physical attributes required to embody a solution principle can be embodied into a physical concept. For example, a cantilever beam with an area facing the direction of sound can be used to embody the physical effect of changing a pressure into a force and that to a deflection, a plate fixed in space parallel to the direction of deflection can form the other plate of the condenser, etc. This means that if various physical effects used within a solution principle can be linked with the corresponding physical structures used in the concepts that embody the effects, these structures can also be strung together, in a way similar to the generation of causal networks above, to generate physical concepts to embody a given solution principle.

• A physical concept intended to embody a specific set of physical effects can embody additional physical effects which may or may not be useful for the intended functionality. Also, due to the specific attributes of the environment (such as the existence of air in the air gap producing a little air cushion between the plates) in which a physical concept is to be used, other un-intended physical effects could be triggered. These additional, un-intended effects are often those which cause the specific problems associated with embodying a solution principle. It is possible to support on computers the detection of these additional effects if physical structures of a concept and the environment in which it is to operate can be represented, and potential physical effects can be allowed to be triggered by the attributes of these structures.

• There are different design problems which can be solved using (parts or combinations of) similar physical effects and solution principles. For instance, parts of the solution principle for probing surfaces are the same as those for acoustic sensors. If these physical effects, and their physical concepts, can be sieved, sorted and stored in appropriate ways, these can be used interchangeably for solving not just one problem in alternative ways, but a whole class of similar problems.

• It is important to realise that some of the physical effects are scale-independent (ie, they apply equally well in macro as well as in micro-scales), some are easily applicable in macro-scales and not in micro-scales, and vice-versa. Therefore, in a similar fashion to the development of generic and application-specific databases for materials [Cebon and Ashby, 1992], generic and scale-specific databases of physical effects can be developed and used within specific scale-contexts.

2 A PROPOSED CONCEPTUAL DESIGN FRAMEWORK

Based on the above observations, a conceptual design framework is proposed [Fig. 1]. Within the framework, it should be possible for a designer to describe the design problem in terms of suitable required functionality, generate possible solution principles, embody them into suitable physical concepts, evaluate these concepts to detect possible unwanted effects, and modify either the concept, the principle, or the functionality to minimise or eliminate these problems.

It seems that the following framework functionalities need to be supported.

• The first step is to allow designer to express the design problem in a way that is understandable and manipulatable by the computer.

• The second step is to support generation of a set of plausible solution principles to a given design problem. Each of these principles could be described as a causal network of a set of physical effects.

• The third step is to support generation a set of plausible physical concepts to embody each of these principles.

• The fourth step is to support envisionment of potential physical effects at work to 'discover' additional unwanted or harmful physical effects at work.

• It should be possible for designer to modify alternative physical effects or solution principles, or even to change the functionality of the problem and go through the whole process all over again.

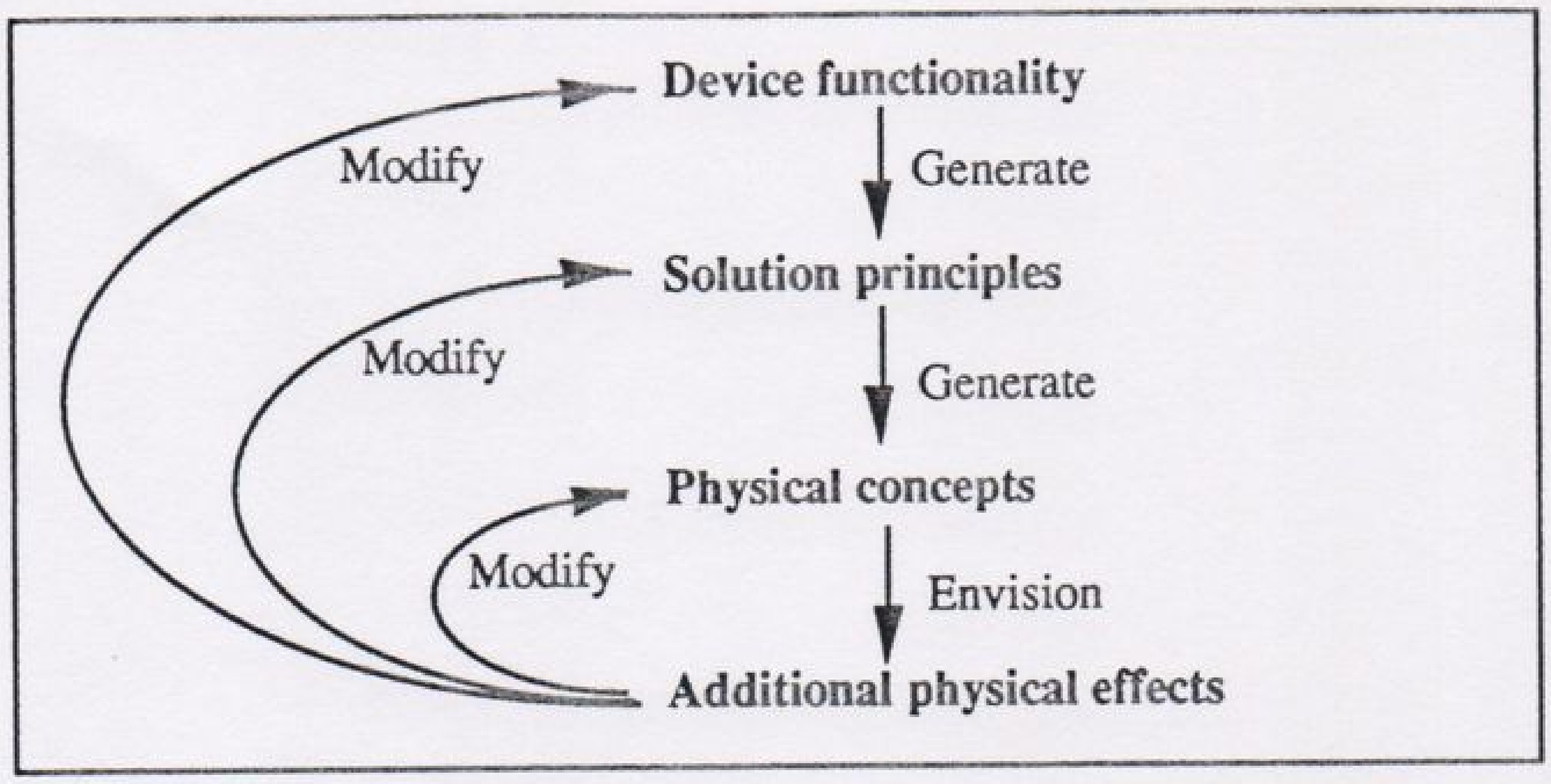

Fig. 1 Proposed Conceptual Design Framework

3 Reasoning, Data Structure and Available Expertise

We need to support the following reasoning :

- Generation of solution principles
- Generation of physical concepts
- Envisionment of physical effects
- Modification of physical concepts
- Modification of solution principles

We certainly need to represent the following information :

- Device functionality
- Solution principles as (combinations of) physical effects
- Physical concepts as (combinations of) physical structures
- Physical processes
- Environment
- Appropriate relationships between these pieces of information

As a starting point, generation aspects can be dealt with using the already existing expertise in synthesis methods within the EDC and RACE [see Chakrabarti, 1991; Taura and Yoshikawa, 1991], and using the knowledge-bases developed within SysFund [Tomiyama et al., 1993]. Similarly, expertise and facilities available with SysFund can be used as a starting point for working on the envisionment aspects. Joint effort is required to work on developing the ontology required. Modification aspects, as a starting point, can be left with the designer, and if it is possible to record as cases the modification steps (s)he takes during conceptual design, these could be used to enhance understanding of these processes with the aim of supporting them in future.

4 RESEARCH METHOD

There is a vast amount of information available about the history of development of various sensor designs (both in macro and micro sensors). This can be a useful starting point for the development and testing of reprsentations and reasoning procedures, and the proposed framework. The research method proposed here is shown in Fig. 2. The idea is to analyse each case history (for instance, the development of acoustic sensors) and find out what is required to be represented and reasoned about in order to support the development process described in the given case history. Once this information is extracted and represented in the framework, it can be tested by testing (i) whether it is possible for a designer to use the framework to go through the major events in the case history (ie, whether the design problem can be described, whether solution principles can be generated, whether these can be converted into the physical concepts described in the case history, whether these physical concepts can be tested for triggering unwanted physical effects, especially those mentioned in the case history, and whether the designer could go through the modification loops of the case history), and (ii) whether using the framework gives any additional benefit. If adequate representation and reasoning procedures, and a knowledge base of resuable functions, physical effects and physical structures could be developed for efficiently supporting a wide range of case histories, this could be regarded as an initial confirmation of the power and potential of the framework.

The rest of this article focuses on the existing work on a brief overview of the EDC and RACE methods for the generation of conceptual designs on computers, and then provides a more detailed account of the EDC approach, with an emphasis on how this can be applied to micro-mechanisms.

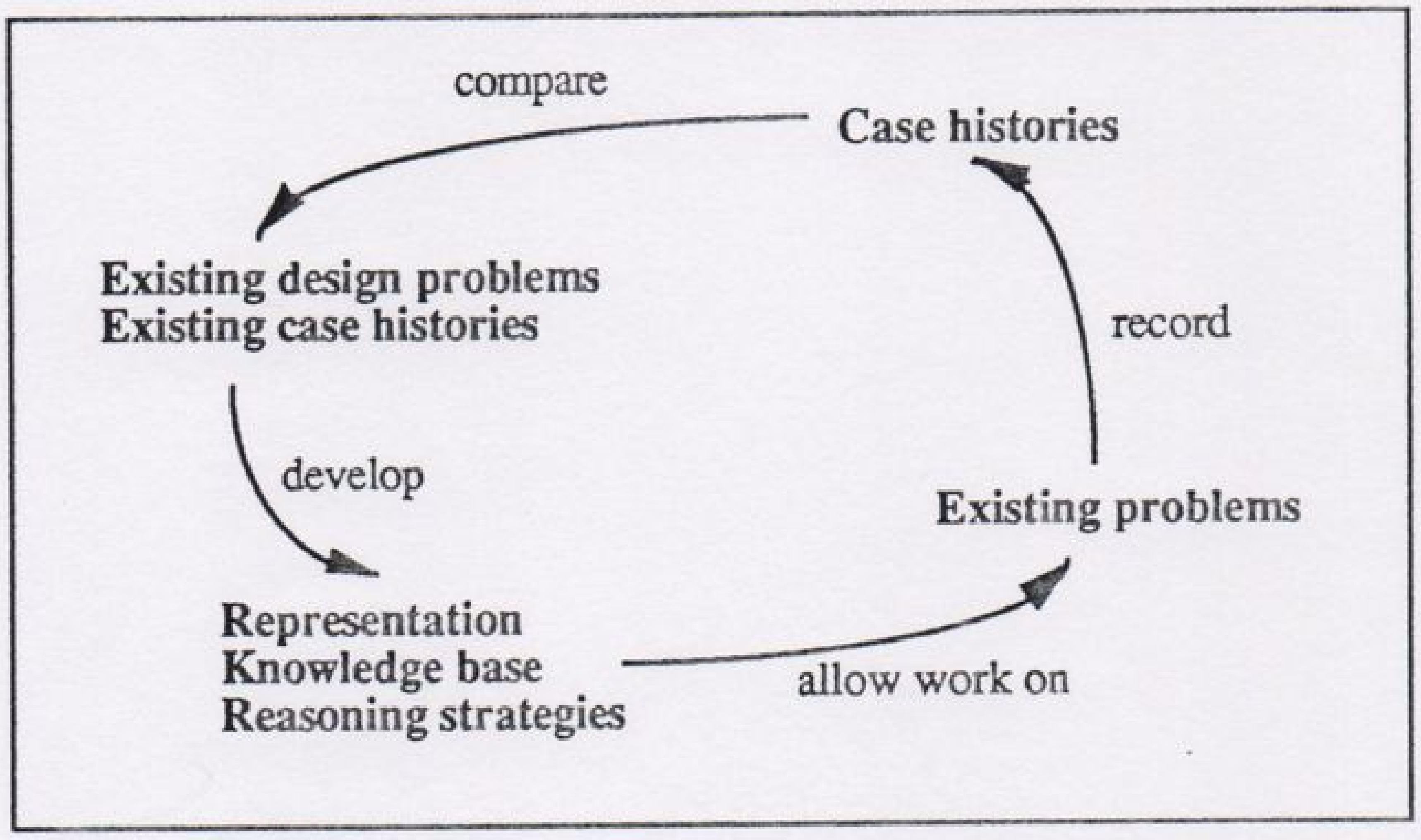

Fig. 2 Proposed Research Method

5 CONCEPT GENERATION METHODS: EDC AND RACE PERSPECTIVES

Both the EDC approach (Chakrabarti, 1991) and the RACE approach (Taura and Yoshikawa, 1991) have been originally developed to support generation of conceptual design solutions to mechanical motion transmission problems. Fig. 3 shows a schematic of the two approaches.

In the development phase of the EDC approach, a set of existing mechanisms were analysed first in order to break down the structure of each mechanism into its motion transmission components (such as gears, shafts, etc). By analysing all the components so extracted, a language, consisting of a set of basic motion transmission elements, with their topological and spatial characteristics, and their combination rules, has been developed. Using this language it is possible to represent the structure and layouts of each of the above mechanisms. The reasoning procedures in the program use the above combination rules using 'search' and 'constraint propagation' techniques to generate conceptual solutions (as input-output networks of these basic elements) to design problems expressed in terms of any number of inputs and outputs having various motion characteristics (such as whether they are forces or torques, what their spatial directions are, etc). All solutions possible using a given the database are generated by the program, as well as alternative spatial layouts for each of these solutions.

In the RACE method, a set of existing mechanisms are stored in the database as combinations of simpler components. The reasoning procedures are based on Genetic Algorithms paradigm; they use existing mechanisms closest to solving the given (single input-output) design problem as initial 'parents', and 'evolve' them (using the principles of Darwinian evolution, such as 'reproduction',' crossover' and 'mutation' between the parts of these mechanisms) in order to end up with one or a few 'evolved' mechanisms as satisfactory solutions [for more details, see Taura and Kubota, 1993].

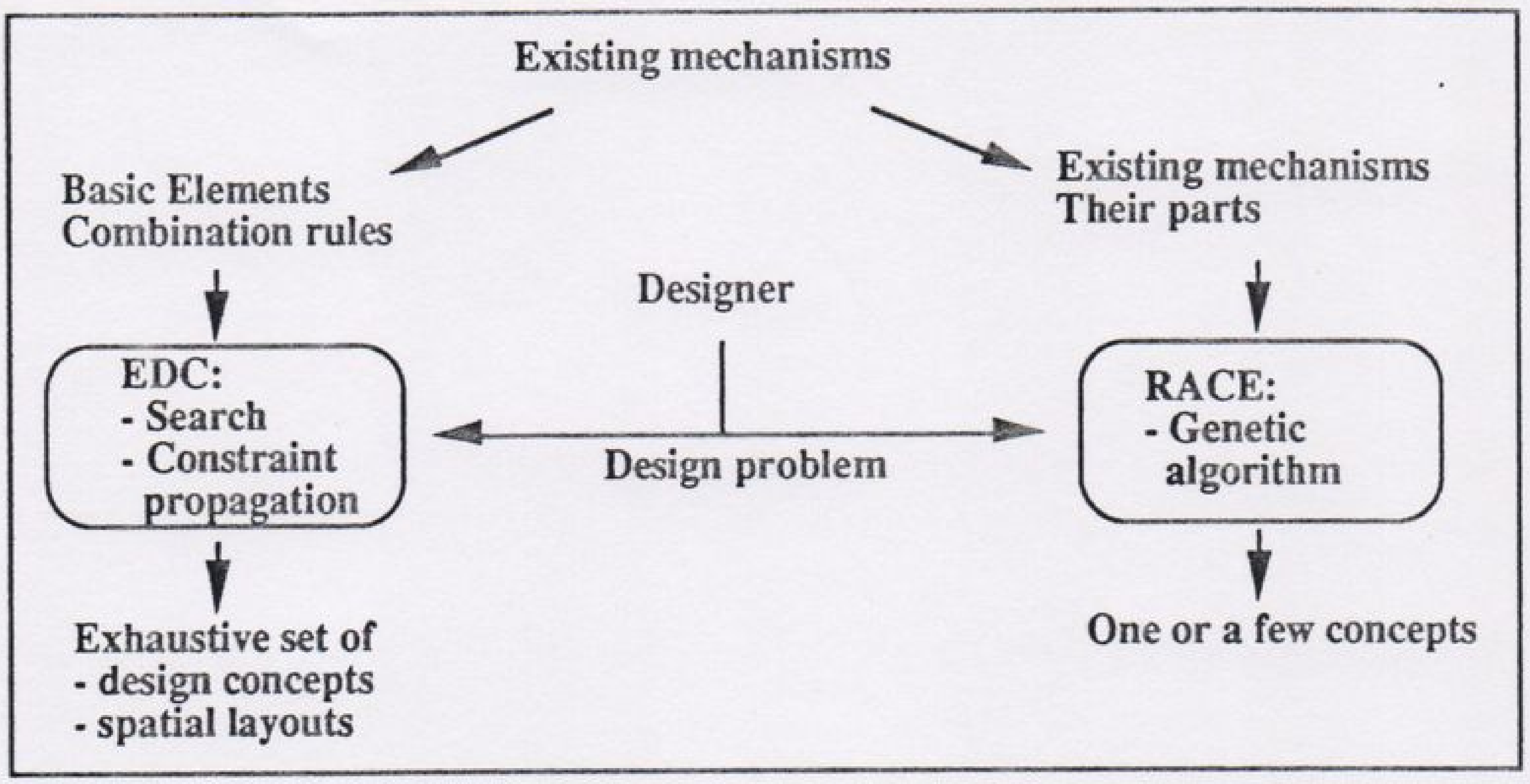

Fig. 3 EDC and RACE Approaches to Concept Generation

6 EDC METHOD TO CONCEPT GENERATION: AN EXTENDED ACCOUNT

The EDC approach is based on the assumption that a design problem, which can be represented using a set of input and output characteristics which can change with time, can be solved by (i) first generating conceptual solutions which solve the problem at one instant of time, and then (ii) using temporal reasoning to check and ensure that it solves the problem at all other instants. The first step, instantaneous synthesis, is presented below. Temporal reasoning is presently under development. For more details on the overall framework, see [Chakrabarti and Bligh, 1994].

6.1 Representation of Design Problems

An instantaneous Multiple Input Multiple Output (MIMO) design problem can be viewed as a transformation between the characteristics of a set of instantaneous input vectors and output vectors (of which Single Input Single Output (SISO), Single Input Multiple Output (SIMO) and Multiple Input Single Output (MISO) systems are special cases). A vector would have a kind, an orientation in space, a sense of its orientation, a magnitude and a position in space associated with it (Fig. 4). The constructs for representing a MIMO design problem involving a transformation between m inputs and n outputs are:

Problem Inputs, Outputs
Inputs $x_1, x_2, x_3 ... x_m$
Outputs $x_{m+1}, x_{m+2}, x_{m+3} ... x_n$
x_i (1..m+n) Kind: (*force / torque / linear motion / angular motion*)
Orientation: (***i*** / ***j*** / ***k***) [*i/j/k* are unit vectors in a Cartesian co-ordinate system]
Sense: (+ / -)
Magnitude: (some number)
Position: $(x_1 \mathbf{i} + y_1 \mathbf{j} + z_1 \mathbf{k})$

6.2 Representation of Solutions

As the input/output are vectors having specific characteristics, the solution structures are vector transformers which transform a set of input vectors into a set of output vectors. The I/O points of the vectors associated with a transformer specify their positions in space, and the spatial separation between these I/O points becomes the position transformation by the transformer. The I/O vectors of a structure are related by various physical principles, which determine the relative orientations, senses, and magnitudes of the vectors involved. All these lead to a representation in which a vector transformer is represented by a 3-tuple of vectors, i.e., an input vector (I-vector), an output vector (O-vector), and a length vector (L-vector). The length vector is defined as a vector joining the input point to the output point, and is directed from the input point to the output point. This vector is created to explicitly reason about the position changes involved in a solution. An I-vector or an O-vector has a kind, orientation, sense, magnitude and position, while an L-vector has a position and orientation (given by the position of the I-vector, and the line joining the positions of the I- and O-vectors), sense (directed from the input point towards the output point), and magnitude (spatial separation between the input point and the output point). These characteristics are variously coupled, depending on the characteristics of the specific transformer involved.

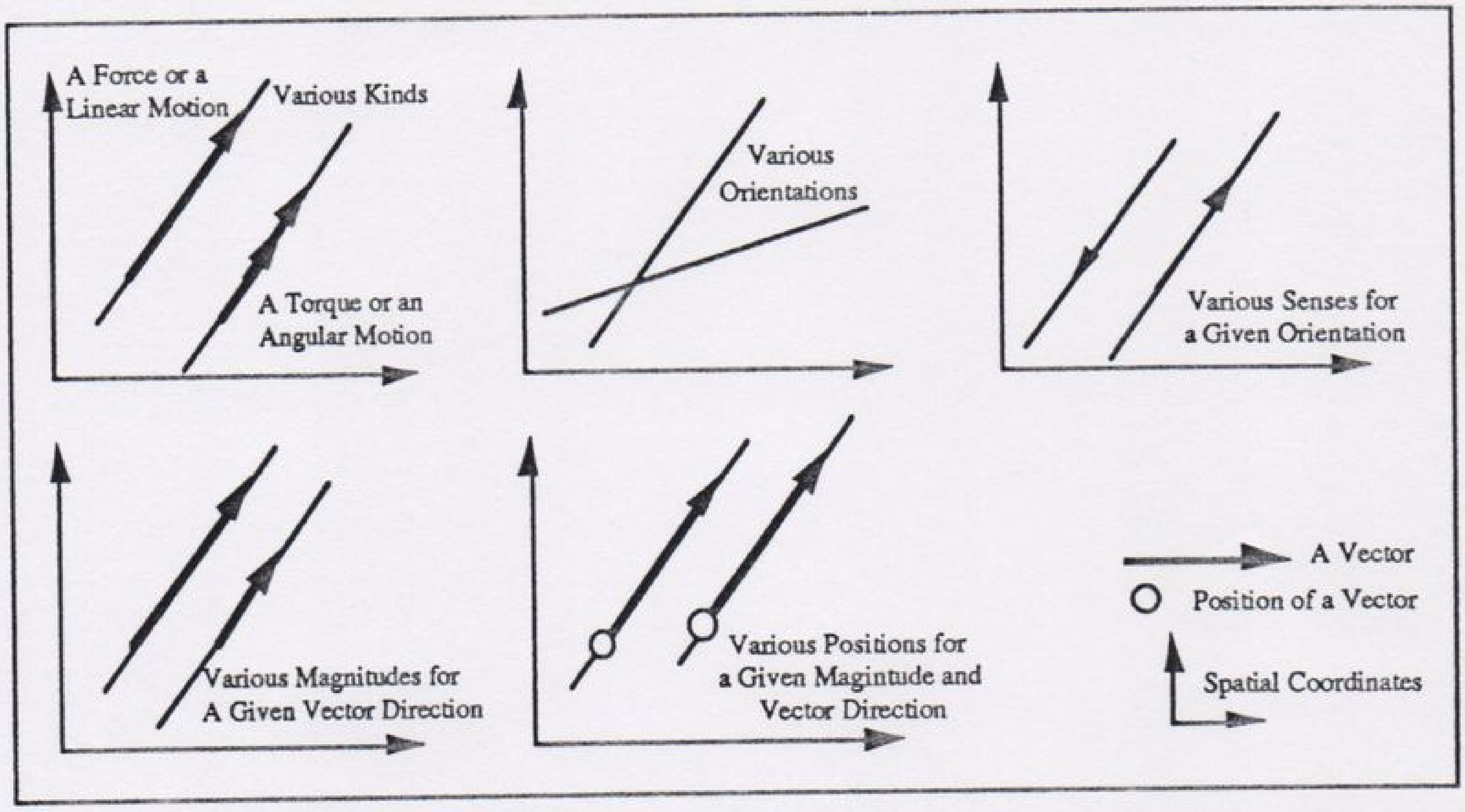

Fig. 4 Various characteristics associated with an I/O vector in mechanical transmission design.

The spatial relation between two vectors can be expressed by using a combination of two properties: whether the lines of their action are parallel (P) , and whether their lines of action intersect (I). Using the combination of these two properties (and their negatives), we find the four possible spatial relations, namely, parallel and intersecting (PI), parallel and non-intersecting (PNI), non-parallel and intersecting (NPI), and, non-parallel and non-intersecting (NPNI). Using the spatial relations among its vector characteristics, a known structure can be typified into one, or a combination, of the types shown in Fig. 5. So, a shaft would be a structure of type PI (Fig. 5) having torque (and/or angular motion) as the input and output kinds which are parallel and intersecting. A solution is defined as a network of structures defined using the above representation, and its I/O properties as well as spatial layouts can be computed from the characteristics of the individual structures of which it is composed.

6.3 Synthesis Procedure

The synthesis procedure takes a database of structures, such as shaft, lever, wedge etc., represented as described in section 6.2, and a specification of a problem in terms of its instantaneous function using the representation in section 6.1, and use search and constraint propagation techniques to generate an exhaustive set of conceptual solution layouts to solve the problem. Take the problem of locking a toilet door and indicating that it is locked. One way of achieving this would be to move a part attached to the door into a slot of the static frame. One way of doing this would be to generate an output translation causing movement into the slot, and another output movement to change a display for indication, from some specified input (such as a couple). One way to represent this function is in terms of two equal and opposite non-collinear forces as input (see I1 & I2 in Fig. 6) and two variously placed motions as output (O1 to represent movement of the part of door into slot, and O2 for display, see Fig. 6). This function, and one of its solutions generated by the program, are shown in Fig. 8. Lever-1 and Lever-2 takes I1 and I2 respectively as input and produces a rotation at their connection. This rotation is converted into a translation by Lever-4 to change display at O2, and into another translation, by a combination of Lever-3 and a Tie-rod-1, to provide the required movement of the part attached to the door into the slot. The given schematic is one of several alternative layouts, generated by the program for this solution, which can in principle satisfy the directions required of the inputs and outputs.

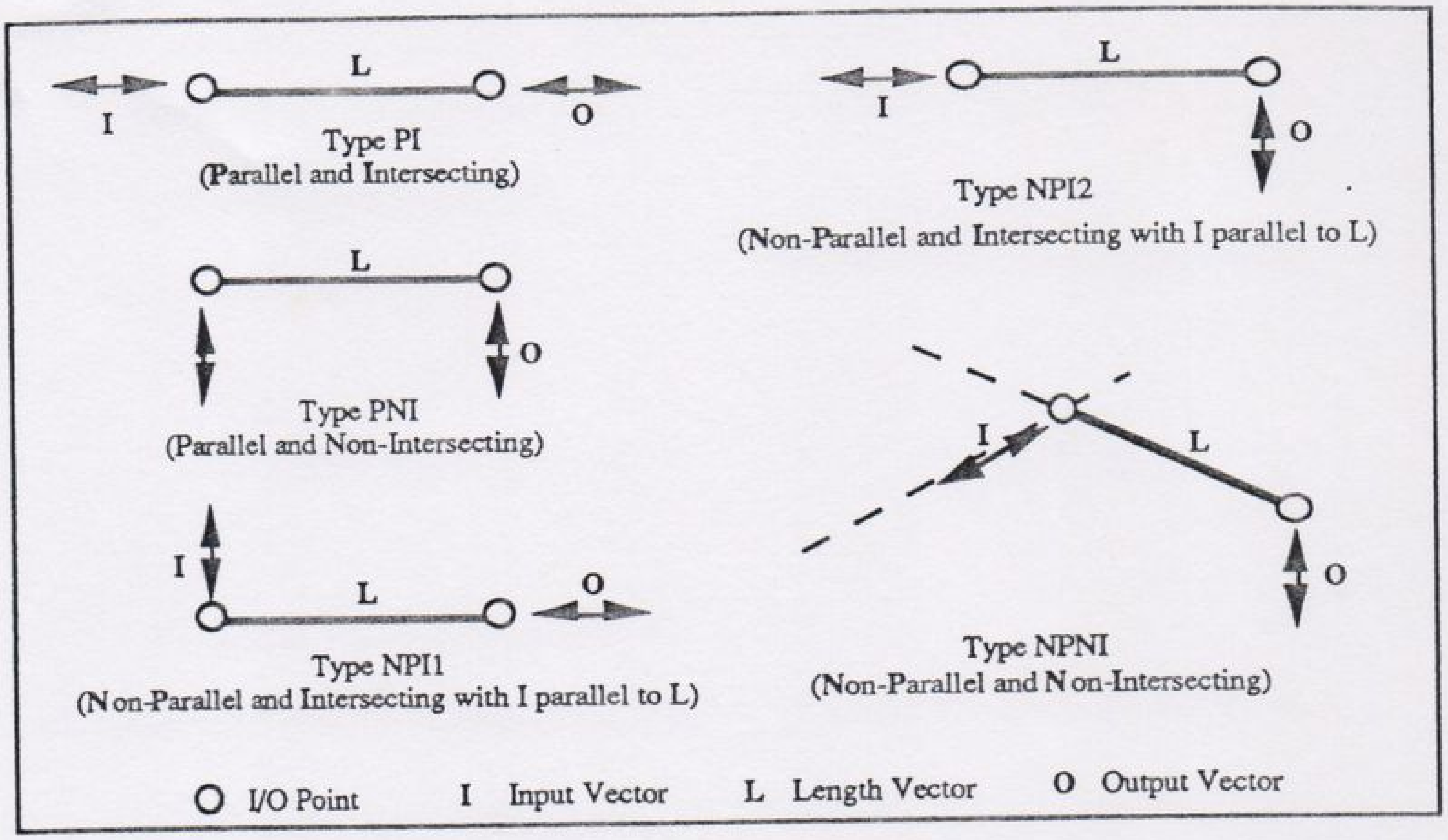

Fig. 5 Various possible spatial relations between the I/O vectors of a transformer.

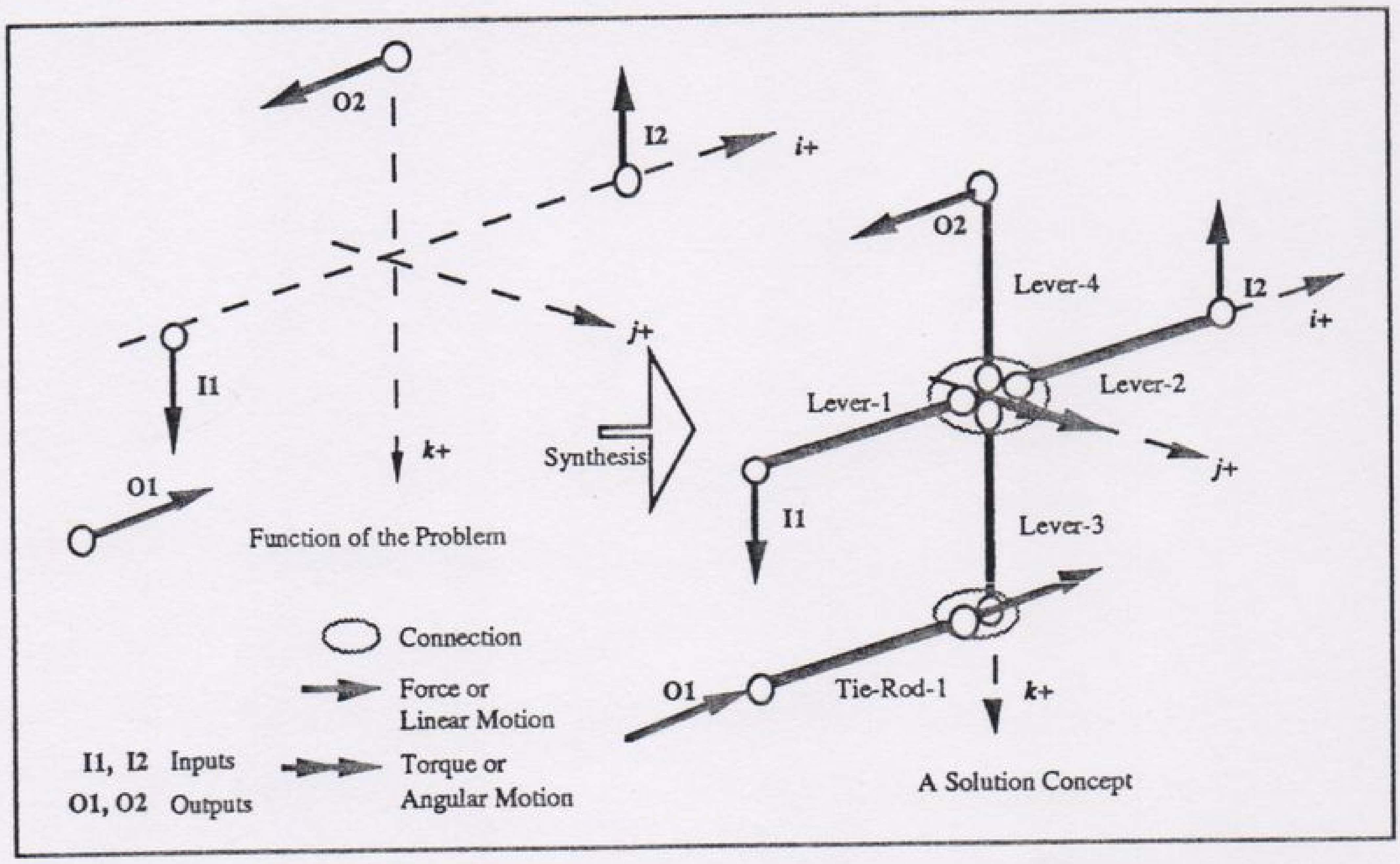

Fig. 6 An instantaneous function of the door-latch problem and one of its solutions

7 MICRO-SENSORS APPLLICATION: GENERATION OF CAUSAL NETWORKS

It is important to note that although the above synthesis procedures have so far been applied to mechanical motion transmission systems, they have wider applicability. Anything that can be expressed in terms of some inputs and outputs, can be used as a 'basic element' for being used by these synthesis procedures. In order to generate alternative causal networks to solve given sensor design problems, we need to define the design problem and the various possible physical effects in terms of their inputs and outputs. By analysing the available sensor designs, these inputs and outputs could be found. For instance, in the condenser principle, the design problem could be expressed as a transformation between an input pressure and any measurable output (such as a change in voltage, in current, etc); the various physical effects can be: (i) a change of pressure as input is transformed into a change of force via an area, (ii) a change of force as input can be transformed into a deflection via a spring stiffness, (iii) an input deflection and a static parallel plate amounts to a change of air gap between them, etc. Once these and other physical effects are extracted, from other solution principles for the same problem as well as for other problems, these can form a bagful of physical effects which can be used for synthesis of causal networks as alternative solution principles. Two such alternative networks for the pressure sensing problem are shown in Fig. 7.

SysFund can provide a platform for the development of the required databases of physical effects. If each physical effect in the database also contains information regarding its scale-applicability, this information can be used to grade the solution principles generated during the synthesis process in terms of their degree of suitability to the given scale. Also, the constraints between the inputs and outputs, such as area, stiffness etc., can be used to develop alternative physical concepts which could embody these physical effects.

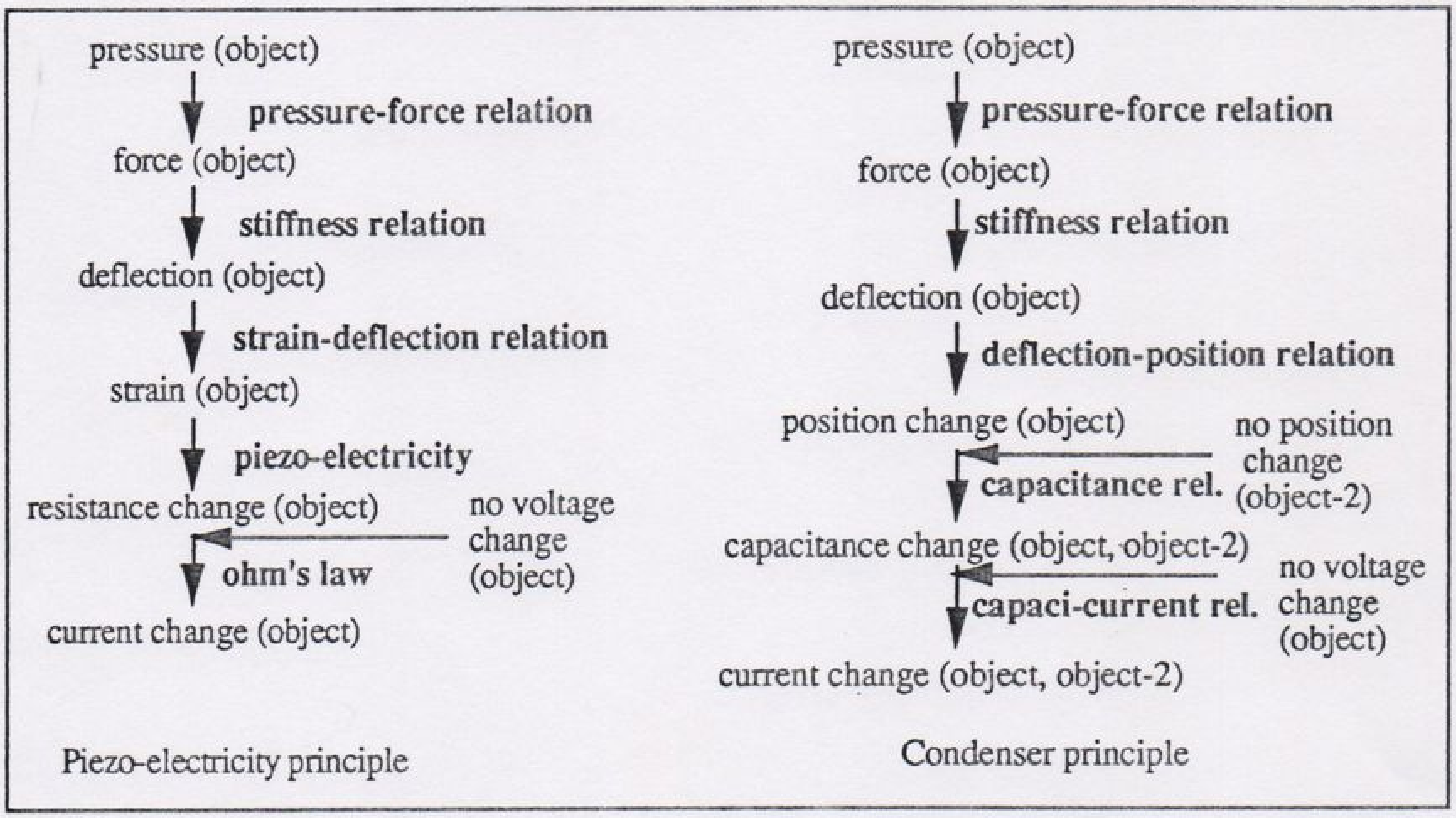

Fig. 7 Two Alternative Solution Principles for Sensing pressure as Electrical Signals

REFERENCES

Sessler, G.M. *Silicon Microphones*, 127th Meeting on the Acoustical Society of America, MIT, Cambridge, MA, 1994.

Chakrabarti, A. Designing by functions, PhD thesis, Cambridge University Eng. Dept., 1991.

Chakrabarti, A., and Bligh, T.P. *A two-step approach to conceptual design of mechanical devices*, AI in Design'94 (J.S. Gero & F. Sudweeks, Eds), Kluwer Academic Pub., 21-38, 1994.

Taura, T., and Yoshikawa, H. The generation of design solutions using metric space methods, Proc. IFIP Workshop on Intelligent CAD III, Yoshikawa, Arbab and Tomiyama (Eds.), Elsevier Science Publishers, 273-287, 1991.

Kubota, A., and Taura, T. A conceptual design process model based on functional similarity, JSME-ASME Joint Workshop on Design, 149-154, 1993.

Cebon, D., and Ashby, MF. Computer-aided materials selection for mechanical design, Metals Mater., 8, 25-30, 1992.

Tomiyama, T., Yoshikawa, H., and Kiriyama, T. Conceptual design of mechanisms: a qualitative physics approach, Concurrent Eng: Automation, Tools and Tech., Kusiak (Ed.), 131-151, 1993.

Brugger, J., Jaecklin, V.P., Linder, C., Blanc, N., Indermühle, P.F., and de Rooij, N.F. *Microfabricated tools for nanoscience*, J. Micromech. Microeng., 3, 161-167, 1993.