Knowledge-based technology in the service of health

Bruce Sayers

The significance of numerical information for decision-making usually has to be assessed on the basis of knowledge that does not readily lend itself to quantification. Computational logic can help to handle the complexity of this process of interpretation without sacrificing the necessary rigour. When this approach is used, health indicators are expressed in the form of a “knowledge map”.

Information is the raw material of human decision-making. In the context of health, decision-making needs information about health status and some understanding of, for example, the origins of health deficits. Information is usually acquired in the form of quantitative data, while understanding is sought from the results of analysis of these data. But this approach neglects a different source of information: knowledge in the form of observations, facts, beliefs and insights that can be gained by expert observers, and which can be a potent resource in understanding important aspects of health in a community. On the other hand, it is not always easy to see how to combine this resource with the formal methodology of public health research which usually employs quantitative data and, very properly, makes rigorous statistically based analyses of these data.

However, this formal methodology needs extension in any event, because it is subject to two difficulties. First, sufficient and reliable quantitative data may be hard to obtain; there is a widespread need for corrections and estimates. Second, research is sometimes faced with entities for which the concept of measurement has no meaning. The classic response to the latter situation is to devise an indicator that offers some kind of, perhaps remote, quantitative representation of the information sought. If the information is itself complex, the indicator may only offer a rudimentary representation of it; this leads to correspondingly limited understanding.

For example, the understanding and use of hygiene in food storage and preparation in a community is a health-relevant aspect of domestic life but not quantitative. It could be seen as a necessity, and schooling might consequently include instruction on such matters, but it does not in itself involve measurement. However, it might be possible to devise an operational indicator linked to some relevant information, such as: “the proportion of women having received instruction on this topic”. This could be useful but there are important questions to which such an indicator is
insensitive. Did the women understand? Did they take the matter seriously? Was there peer group or other pressure to disregard it? Were there cultural reasons for them not to bother? In short, measurement is not really appropriate for expressing “knowledge” about the actual situation, but any competent observer of the community could see and verbalize the realities. Among other things, an expert might note, for example, that “in most households, food hygiene is not considered and not practised”. Such an observer might add: “In this community, older women in the household hold the power. They do not understand or accept notions of hygiene in relation to food”. This might be sufficient to explain the factual observation and could be relevant to any intervention being considered. Such verbalizations constitute “expert knowledge” in the sense implied here.

This kind of example illustrates the scope for a new approach, that will supplement the classic formal methodology by using expert knowledge in some systematic way.

Modern computer science is much concerned with “natural language” processing and, specifically, with the analysis of knowledge in the form of verbal statements.

The basis of the formal methodology is the idea that the totality of what is meant by the word “health” can be adequately captured in numbers. The idea is linked to a very proper desire for objectivity and rigour, but these attributes are no longer restricted to quantitatively based investigations. Modern computer science is much concerned with “natural language” processing and, specifically, using computational logic for the objective and rigorous analysis of semantic information, or knowledge in the form of verbal statements. Since this is the form in which expert observers express their observations and insights, it should offer new ways to capitalize on expert knowledge as a resource. At the same time, non-quantitative aspects of health can be more directly incorporated into the description and evaluation of health. Developments like these are sufficiently important to warrant the investment of considerable effort; they open up significant new pathways to understanding the problems of global health.

Knowledge-based health indicators

The aim of this approach to using knowledge is to design new types of health indicator to replace quantitative indicators that need data which cannot be obtained, or to supplement existing indicators by handling health-related factors that cannot be quantified at all. If this is successful, the same methodology could be applied to the study of intersectoral interactions involving health. The relevant knowledge would be obtained by using expert observations and insights in the form of verbalized observations, facts, beliefs and insights derived from the study of specific communities and analysed for inherent consistency. The methodology would use computational logic to test consistency, make inferences and explore implications.

Health being a complex concept, the word means many things to us simultaneously. It is by no means related exclusively to disease. It can refer to the presence or absence of disability, the effects of behaviour patterns, the adequacy of nutrition and education, various kinds of environmental circumstances, availability of safe
water and sewage disposal, accessibility of medical care, quality of housing, the existence of cultural factors that affect health, and other aspects of human well-being. These are all elements contributing to our idea of health, yet few of them can be measured. They are, in effect, more or less universally accepted concepts.

Health status as a concept is thus formed from a number of elements, some or most of which are themselves concepts. When each is examined in turn for what it contributes, many of these contributions will be found to be concepts as well. In principle, this process could be continued until, finally, it is found that the contributory elements can be observed or measured. If perceptions and knowledge are adequate and accurate, they can provide the starting point for making a new picture of health status. The totality of the linkages and the knowledge expressing how elements are related constitutes a knowledge map of the indicators of health status. It shows how an assessment of the health status of a given community is assembled from the observations of basic facts in the light of other insights about relationships. The building blocks of this knowledge map would be the pieces of semantic or other information that express how the inputs at one level interact to contribute to the relevant concept at the next level up. Such a knowledge map is the target of the research effort.

In practice, the present design strategy uses a “bottom-up” approach. It starts from the information about the community that comes from expert observers. This constitutes the input information. It is supplemented with insights about the significance of these observations and the inferences that can be made from them again by the same expert observers. These preliminary inferences constitute a foundation on which the computational logic process can build the knowledge map. The construction of the knowledge map then proceeds by exploring the inferences and the various interactions between the observations. In practice, the array of available information and the complex network of deductions and inferences to which they give rise are too elaborate for pencil-and-paper analysis. This is why the process must be automated, using logic programming to formalize it. The knowledge map thus receives inputs from expert observers, who also contribute insights about at least some of the implications of the observations, and these are used as starting-point inferences. The output is the knowledge map, which constitutes a health status indicator.

**Example: health status of the elderly**

The process can be illustrated by investigating a specific indicator: “health status of the elderly in a community”, devised for the purpose of explanation. Experts could structure their observations in various domains of observation. For instance, one domain may be prior occupational exposure to a potentially health-damaging environment in relation to the immediate circumstances of the elderly. Another may be the economic position of the household. Another may be the relevance of family structure to the health status of elderly members. Other domains may concern the nutritional position, physical
and mental health of the elderly and the extent of any community support. Additionally, expert observers of a community would have general background knowledge pertaining to the demographic and social structure of the community, its cultural setting and its behaviour patterns. Similarly, a large amount of general knowledge about health is also available. Each of these domains constitutes a knowledge base which needs to be drawn upon in developing a knowledge map describing the community.

Inferences follow from the combination of different pieces of knowledge derived from the various domains. They can be expressed in the form of logical rules which indicate the consequences if the stated conditions apply. Expressed in this way, these rules can be used to test other sets of input observations that would pertain to, for instance, other types of household in the community being examined. The following examples will help to show how the process operates.

Since the household contains school-aged children and school-aged children are susceptible to respiratory infections, such infections will be repeatedly introduced into the household. Because of pulmonary weakness caused by long-term occupational exposure to stone dust, exposure to infectious children will result in repeated respiratory infections in the elderly person. Repeated episodes of infection lead to progressive deterioration of pulmonary function. In the presence of nutritional deficiencies which accelerate the process, this results in progressive limitation of physical activity and, consequently, increased physical dependence. Since cognitive capacity has not significantly changed, increased physical dependence, coupled with an inability to contribute to household income or to assist in household tasks, often leads to mental depression.

These are some of the deductions which may be formalized into logical rules leading to logical outcomes. Certain of the rules involve direct deductions from the inputs, while others use the deductions from other rules to take the process further. All these initial rules can be generated by the experts from knowledge of the community, drawing also upon background information.

It may be possible for the experts to produce further rules, but when other domains are taken into account, the task rapidly becomes too complex to handle without automated assistance. This is particularly true when there are different options for the inputs, leading to different outputs. A direct statement about the health of the elderly, assuming for the moment that all households were identical, would consist only of the combination of observations on physical and mental status, whereas the knowledge map interprets the observations. It only allows physical status to be viewed through the effect of rules which include observations about the likelihood of repeated episodes of respiratory infection, of the progressive deterioration of pulmonary function and physical capacity leading to physical dependence. Mental status is also viewed through the effect of rules which include
other relevant observations. The application of these rules to simple observations therefore produces the information needed for this (rudimentary) knowledge-based health indicator.

Four points about the indicator can be noted.

- It is multidimensional, in that the outputs of the rules are of different kinds and cannot usefully be added up in any simple way.
- It is possible to track backwards through the knowledge map to find out how the final result was assembled, or what caused it.
- Each type of household encountered would contribute proportionally to the overall indicator.
- The indicator output in each dimension could initially make use of qualitative grading, such as low, medium or high risk.

In the form outlined, the knowledge map does not yet make significant use of the inferential power of computational logic. From this point it should be possible to make and explore automatically the further inferences that are feasible. But the above example should be sufficient to show how expert knowledge can be used, and to clarify the concept of “knowledge-based” indicators.

**Priority setting and resource allocation**

The desirability of taking a systematic approach to priority setting is now recognized. The same is true for resource allocation, which is a more general version of the same activity. However, it is desirable to translate this principle into a practical decision-support methodology that will ensure objective and transparent recommendations about where research effort should be placed and where resources should be allocated to get the desired results.

Setting priorities depends upon testing research options against criteria, although an acceptable list of criteria for priority setting in health research still needs to be agreed. One relevant criterion can be expressed as “the scale and urgency of the need, based on health level assessment, and on an understanding of the fundamental determinants of the health problem”. But there are many others, such as “the availability of human and other resources to do the work”, for example, or “the existence of non-health sector options for intervention”.

Whenever resources have to be allocated, modern planning technology may help. This can be illustrated through an industrial example. Consider the case of a company involved in many simultaneous civil engineering construction projects. Each project has its own needs for personnel, equipment and raw materials which demand the allocation of appropriate resources in the most cost-effective way for the company. But such tasks – for instance, allocating heavy construction equipment to a large number of simultaneous construction projects – are constrained by many factors. These include: the time
periods during which the project contracts must run, the different equipment needed for the various projects, the different equipment owned by the company, the mandatory maintenance schedules for the individual pieces of equipment, the limitations due to the number of maintenance personnel and maintenance workshop space, and the possibility and costs of leasing particular equipment. Allocating equipment effectively in the face of these constraints is indispensable for the company’s survival and success.

Correspondingly, in planning health interventions, one’s freedom to select appropriate and useful measures is frequently circumscribed by limitations of many kinds, such as acceptability, cultural factors, costs and practicability, all of which can also be called “constraints”. In fact, constraints dominate the task, and make some form of guidance necessary for effective action. A powerful development of computational logic known as “constraint logic” is now available, and is starting to be used for large-scale industrial and commercial purposes where hard-headed decision-making is the rule.

The strategy of constraint logic programming is to allow for and search through all the available allocation options, discarding those that are unworkable because of the constraints they would encounter. Constraint programming means devising technological processes for identifying the logical structure of the relevant constraints automatically, and then using these processes to set up an efficient system for determining the acceptability of options. This turns the impossibly large set of all possible options into a manageable set of possibilities that can then be reviewed for effectiveness and benefit.

For global health, constraint logic programming offers a potentially powerful approach to setting priorities and allocating resources. More generally, knowledge map methodology based on expert observation has a range of applications, such as: indicators; health effects of, say, environmental pollution; and analysis of intersectoral interactions.