DOCUMENT RESUME

ED 309 113 SO 020 022

AUTHOR Sneed, Joseph D.
TITLE Micro-Economic Models of Problem Choice in Basic Science.
PUB DATE 15 Aug 87
NOTE 19p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Evaluation Criteria; Input Output Analysis; Needs Assessment; *Research Administration; *Research Design; Researchers; *Resource Allocation; Resources; Sciences; *Scientific Research; Systems Analysis; Values
IDENTIFIERS *Microeconomics

ABSTRACT

Certain representations of basic scientific knowledge can be coupled with traditional micro-economic analysis to provide an analysis of rational research planning or agenda setting in basic science. Research planning is conceived of as a resource allocation decision in which resources are being allocated to activities directed toward the solution of basic scientific problems. A structuralist representation of scientific knowledge will be employed to provide a relatively precise characterization of a basic scientific problem. The main thrust of the analysis consists in describing various ways in which values enter into those decisions. Distinction is made between internal and external values operating in these decisions. The discussion focuses on exploiting this representation of knowledge in a manner which will illuminate the scientific value of solving certain basic problems. A micro-economic approach to the problem of resource allocation suggests that it may be fully understood in terms of values (or utilities) and probabilities attached to the solution of the problems by the decision makers. A bibliography of 12 items is included. (Author/PPB)

# Micro-Economic Models of Problem Choice

# in Basic Science



Joseph D. Sneed
Department of Humanities and Social Sciences
Colorado School of Mines
Golden, CO
and
Center for Values and Social Policy
University of Colorado
Boulder, CO



8/15/87

## I. Introduction

In this paper I will describe the way in which a certain representation of basic scientific knowledge can be coupled with traditional micro-economic analysis to provide an analysis of "rational" research planning or "agenda setting" in basic science. Research planning will be conceived as a resource allocation decision in which resources are being allocated to activities directed toward the solution of "basic scientific problems." A "structuralist" representation of scientific knowledge will be employed to provide a relatively precise characterization of a basic scientific problem. The main thrust of the analysis will consist in describing the various ways in which values enter into these decisions. In particular, some effort will be devoted to distinguishing "internal" and "external" values operating in these decisions. The discussion will focus on exploiting this representation of scientific knowledge to say something about the "intrinsic scientific value"--"scientific value" for short--of solving certain problems basic scientific problems. This work is a part of a larger project whose aim is to describe the more general and typical situation of research planning in which technological and other external values play a role.

## II. Problems in Basic Science

The fundamental concept I will use to analyze resource allocation in basic science is the concept of a "scientific problem." Roughly, decision makers in basic science--individual scientists, research teams, research administrators--allocate resources to attempts to solve scientific problems. A micro-economic approach to this decision problem suggests that it may be fully understood in terms of values (or utilities) and probabilities attached to the solution of the problems by the decision makers.

My aim here is ultimately to clarify the nature of the values that might adhere to the solution of scientific problems. This aim is, I think, best approached by considering first how one might characterize a basic scientific problem. The account I have to offer of basic scientific problems is derived from a general picture (or representation scheme) for scientific knowledge. This scheme is described in detail in [1].

## III. A Representation of Scientific Knowledge

We may represent basic scientific knowledge generally as a net (directed graph) consisting of "theory elements" (nodes) linked together by intertheoretical links (arcs). Intuitively, theory elements are the smallest units of scientific theory that may be used alone to say something intelligible. They are the elementary building blocks out of which more complicated pieces of scientific theory are constructed. For example, Newton's theory of the gravitational force corresponds to a theory element which is one part of the full "theory" of Newtonian particle mechanics. The links between theory elements are intuitively paths permitting certain kinds of information to be passed between the elementary building blocks of scientific theory. The direction of the link indicates the direction of information flow.

Typically, each theory element T is linked to other theory elements by links going in two directions. That is, T receives information <u>from</u> other theory elements and, in turn, passes information <u>to</u> other theory elements. Links may even go both ways between two theory elements. That is theory elements $T_1$ and $T_2$ may <u>exchange</u> information. For simplicity, I ignore this possibility here.

Theory elements are the intellectual focus of the social activity of scientific problem solving. Scientific communities may be organized around theory elements and/or collections of closely linked, "neighboring" theory elements--"theories" in one common sense of the word. These theories set the problems for the community as well as provide the means and criteria for the solution of these problems. "Theories" in this sense--e.g. Newtonian particle mechanics--play a role somewhat analogous to "paradigms" in Kuhn's account of "normal science" [5]. In the sociological literature on "theory choice" ([6],[12]) it appears that the "theories" that are being chosen are more often to be represents as our theory elements than as "theories" in the larger sense. Choice among theories in the larger sense appears, at least in "mature" areas of science, to be confined to those periods of scientific activity that may be viewed as truly "revolutionary" in Kuhn's original sense. For present purposes, it is necessary to see only a bit more precisely how this works. Those interested in a detailed, formal account may consult [1].

Each theory element T consists of a "conceptual core" K and a range of "intended applications" I. It is the range of intended applications that provides the most familiar kind of scientific problem and, as well, the key to characterizing other kinds. The range of intended applications consists of those problems that scientists committed to using T recognize as "soluble." The basis of this recognition has been the focus of some attention in the sociology of science ([2],[4],[6],[8],[11],[12]). For the moment, it suffices to say that our representation of scientific knowledge locates the considerations relevant to this choice in the configuration of links among theory elements. We will see below more precisely how this works. It is however clear already that our account is consistent with the "methodological dogma" that "...scientists define some problems as pertinent, and others as uninteresting or even illegitimate, primarily on the basis of theoretical commitments and other assumption structures." ([12], p.74).

According to the knowledge representation scheme under consideration, the aim of those practicing science with T is to use K to make some claim about the whole of I which is generally not reducible to a conjunction of claims about individual members of I. This claim is "holistic" in a minimal sense just in that it is irreducibly about all of I. The typical "problem situation" associated with a theory element T = <K,I> is that this claim will have been shown to be true for some finite sub-class of I--call it A. "Open problems" are then members of I~A. That is, an open problem is showing that a member of I~A can be added to A preserving the truth of the theory element's claim for the augmented set. "Problem choice" is then--in the simplest case--just choosing which members of I~A work on.

The intended applications I of a theory element T = <K,I> are determined by the links running to T from other theory elements. Intuitively, these links provide the data about which T is a theory. In the language of traditional logical empiricism, they provide empirical interpretation for some parts of the vocabulary of K. In many cases, other theories T' with links going to T may be viewed as providing means of measuring the values of empirical quantities in K. In these cases, T' may be viewed as a theory about the measurement of these quantities. More generally, these links provide the context for the practice of science with theory element T. They link this practice with other parts of science. Developments in other parts of science--e.g., discoveries that result in new instruments and/or new methods of observing certain things impinge on the practice of T-science via these links. In this way, the links determine the intended applications I that are a part of practicing T-science.

This representation of the intended applications of T will, in many specific cases, do some violence to our ordinary usage of the word 'theory.' The kinds of knowledge that legitimate the results of measurement and observation will frequently be more accurately viewed as "common sense" rather than "theoretical" knowledge. In some cases this common sense knowledge may be viewed as providing the intellectual focus of a technological or experimental tradition. This fact occasions only linguistic abuse, provided that "common sense" knowledge can be represented with the same formal tools as more explicit "theoretical" knowledge. At this time, whether formal representation of common sense knowledge is possible with the tools sketched here is an open question. We consider briefly below the possibility of formalizing the knowledge embodied in technological traditions. See, however [3].

Problem solving with T works in such a way that successfully solved problems in T provide part of the description of the intended applications I" of theories T" = <K",I"> with links going from T to T". Thus, problem solving with T may be seen as a kind of transformation or filter on information. Information coming into T FROM theory elements $T'_1,\ldots,T'_n$ is "filtered" into the results of successfully solved problems and then passed on to other theories $T"_1,\ldots,T"_m$ to partially determine their intended applications $I"_1,\ldots,I"_m$.

To understand the kind of information processing that counts as problem solving with T we must be more precise about the intended applications I and the claim that T makes about them. To do this, I must first say more about what is in the conceptual core K of a theory element. A core K consists of four parts:

$M_p$ - the vocabulary or conceptual apparatus characteristic of the theory element;

M - an empirical law or systematization formulated in the vocabulary stipulated by $M_p$;

it is convenient to think of $M_p$ as the class of all systems that may be described using the vocabulary characteristic of the theory element and M as a sub-class of those structures satisfying some specific law. Both $M_p$ and M are represented formally as classes of "models" in the mathematical logician's sense.

C - some conditions or constraints, analogous to the laws M, that are imposed on collections of systems from $M_p$, rather than on individual members of $M_p$.

Intuitively, K contains means of imposing restrictions both on individual members of $M_p$ and on collections of members of $M_p$. These latter conditions entail that problem solving with K will generally have certain "holistic" features which are important for understanding problem choice. Formally, C may be conceived as a class of sub-classes of $M_p$ satisfying some specific condition on such sub-classes.

$M_{pp}$ - the vocabulary of the theory element that is linked to other theory elements via "interpreting links" going to T--the non-theoretical conceptual apparatus.

Formally, $M_{pp}$ is also conceived as a class of systems that are all "fragments" (or sub-systems) of the systems in $M_p$. Intuitively, these are the systems about which other theory elements may "possibly" provide information via links that "interpret" (or assign meaning to) the concepts that comprise them. 'Possibly' means here that there is a purely formal (conceptual or definitional) link between the concepts in $M_{pp}$ and those in $M_p$ in some other theory. The concepts that appear in members of $M_p$, but not in members of $M_{pp}$ are "uninterpreted" or "theoretical" concepts for the theory element T. They represent concepts that are "internal" to T in that they derive their meaning solely from their use in solving problems associated with T and, perhaps as well, from their role in theory elements that T interprets. This distinction recalls the distinction between "theoretical" and "observational" terms in traditional logical empiricism except that it is explicitly localized to a single theory element.

The intended applications I of a theory element are simple a sub-class of $M_{pp}$. Intuitively, I is that sub-class of $M_{pp}$ consisting of systems for which other theory elements can "in fact" (rather than just as a formal possibility) provide information. That is, members of I must ultimately be linked to "solved" problems in theory elements that interpret T. Other theory elements linked to a theory element T by interpreting links may be thought of as providing "data" about members of I.

Theory element T = <K,I> is "about" I in the sense that it says something about it. The "claim" of theory element T is that all members of I can be filled out with the full (theoretical) vocabulary of $M_p$ in a way that satisfies <u>both</u> laws M and constraints C. Very roughly, the problems for T are the members of I and solving such a problem is providing a specific configuration of theoretical vocabulary that fills out a member of I to member of M. Generally, this problem solving occurs in a situation in which other such problems have already been solved so the filling out is also required to preserve the satisfaction of C when it is added to the stock of already solved problems. More subtle accounts of scientific practice consider the possibility that problems are sometimes deleted from the stock of "solved" problems. We ignore these subtleties here.

The metaphor of problems solving as "filtering and transforming information" may be fleshed-out in the following ways. First, not all possible configurations of information transmitted by links <u>to</u> T may be amenable to "filling out" with theoretical concepts of T in the way just described. In most cases, the claim of T will be non-vacuous in that it will "reject" certain possible configurations of data. In this way, T acts as a filter. But, it also transforms the information it receives from its interpreting links in an essential way. Most obviously, the theoretical concepts that appear in successfully solved problems for T do not appear in any recognizable form in the data that sets the problems. More precisely, they can not be explicitly defined in terms used to describe the this data. Yet it is these full-bodied solutions that T transmits to other theories for which it provides partial descriptions of data. Less obviously, the links themselves may serve to "transform" concepts from one theory into those of another.

We may now see, somewhat more clearly, how the intended applications I in T = <K,I> are determined. I consists just of those members of $M_{pp}$ that are (or can be) linked to successful intended applications I' in theories that are linked to (interpret) T. Minimally, 'successful' means members of I' that have been filled out with theoretical concepts of T' = <K',I'> in the appropriate way. Less minimally, successful applications of T' will have to be linked to successful applications of other theories (if any) besides T that T' may serve to interpret.

In this manner, the net surrounding a theory element T = <K,I> may be seen as providing an intentional description of the intended applications I. It does this recursively by telling us how to check a candidate for membership in I by tracing its connections through the links it has to other intended applications of other theory elements linked with T. This tracing back terminates when it reaches theory elements $T^{\sim}$ in which extensional descriptions of successful applications (members of $A^{\sim}$) are found. Otherwise, at every theory element T in the recursive tracing back a new problem for T is generated.

## IV. Elementary Basic Research Programs

Using this the account of problems for a theory element just sketched, we now turn to the question of choosing the order of attacking these problems. Here too the sociology of science literature offers some guidance, though sketchy. A summary of this work ([12]), p.82) suggests that "...two criteria were most frequently used in selecting from arrays of previously identified problems: (1) the assessed scientific importance of a problem...and (2) the feasibility of arriving at solutions." This summary suggest that the apparatus of micro-economic decision theory might be applicable to providing a more detailed theory about this kind of problem choice. It appears quite natural to identify "assessed scientific importance" as a value or utility and "feasibility of arriving at solutions" as a probability. For the moment we will leave open the question of just whose values and probabilities these might be and focus on the question of using the representation of scientific problem solving just sketched to provide a somewhat more precise formulation of this decision problem.

To begin, let us conceptualize problem choice as a resource allocation problem--one chooses to work on a problem by allocating resources to it. 'Resources' is to be understood intuitively to consist personnel and equipment devoted to the solution of a problem. Let us also conceptualize the problem as one of choosing a "research program"--a plan for allocating resources to problems over some period of time.

We may consider one very simple kind of "research program" for a theory element T = <K,I>. We call these 'elementary research programs' to indicate that they are exclusively concerned with a single theory element. Elementary research programs are very likely an idealization in that no real piece of scientific activity may plausibly be modeled as an elementary research program. Nevertheless, they are a useful starting point for an analysis and may well prove to be fundamental parts of more complex research programs.

Let us suppose that the program begins with some problems for T already solved. That is, at time t there is a finite sub-class $A_t$ of I for which there are at hand extensional descriptions of theoretical emendations At that satisfy M and C. The $A_t$ could be the "paradigm" problem solutions guide the practice of normal science on Kuhn's account, but they need not be. Generally, they just represent the stock of successful applications of T at any time t in the history of the community that practices science with theory element T. We assume that $A_t$ is non-empty and leave the limiting case of an empty $A_t$ for special treatment. For simplicity assume I is finite.

In the situation just described, a resource allocation problem for T-problems may be formulated in this way. How is a finite amount of "resource" L to be distributed over the problems in $I \sim A_t$? At this point we need not be too specific about what L is. We need only to think of it as something measured by a monotone increasing real valued function. More subtle treatments might take it to be characterized by a real vector. For example, one might want to take time as a component of this vector requiring special treatment.

How should we conceptualize a "research program" aimed at solving problems in $I \sim A_t$? Most generally, one might conceive of a research program as a sequence of "allocation vectors"

1) $r(1), r(2), \ldots, r(n), \ldots$

where each vector r(t) represents an allocation of resources to members of I at stage t in the program. That is, r(t) is of the form:

2) $r(t) = \langle r(t,1), r(t,2), \ldots, r(t,m) \rangle$.

The resource vectors are subject to the following constraints.

3) $\sum_{t,i} r(t,i) = L$.

4) $r(t,i) = 0$, for all i in $A_t$.

A somewhat more realistic formulation might replace '=' in 3) with '≤', but little insight appears to be gained by this added complexity. We remain vague about the interpretation of the "stages" in a research program.

If we let 'R' denote the class of all research programs satisfying the above conditions (1 - 4), then "research planning problem," at the scientific community level, may be conceived simply as choosing one among the feasible research programs in R. The one chosen should be among those that are optimal with respect to some value or "objective function" for the scientific community. Our purpose here is to describe this objective function.

We may proceed by introducing some notation to describe the "solution state" of problems in I. Let $Z = \{t_1, \ldots, t_m\}$ be the sequence of stages in the research program, $I = \{i_1, \ldots i_n\}$ some arbitrary ordering of I and

5) $s: Z \rightarrow \{0,1\}^n$.

The value s(t) is to be interpreted thus: problems corresponding to positions where 1's appear in s(t) have been solved and those corresponding to positions where 0's appear have not. Each such function s describes a "possible problem solution history" or "solution sequence" and we may denote the set of all such histories by 'S'. Clearly, we idealize in regarding problems as simply "solved" or "unsolved."

This interpretation suggests that at least one additional condition might plausible be imposed on members of S. The condition that "problems never become unsolved" is may be expressed by:

6) $s(t+1) \geq s(t)$,

where the partial ordering "≥" on vectors means that their respective components are partially ordered by the usual ≥- relation on real numbers. This condition is only plausible on a rather strong interpretation of "solved." On this interpretation, a problem is simply not "solved" if their remains a possibility that a mistake was made which will be subsequently uncovered resulting in the problem's status reverting to "unsolved." With a weaker and more common interpretation of "solved", one would have to countenance the possibility the problems sometimes became "unsolved."

It should be noted that S will include solution sequences in which problems are solved "one-at-a-time." That is, s(t+1) will contain exactly one more '1' than s(t). But it will also contain solution sequences in which more than one problem becomes solved in a given time period. That is, s(t+1) may contain several more '1''s than s(t).

To begin formulating an objective function for an elementary research program, let us first suppose that some value extrinsic to theory element T is assigned to each problem solution vector which is represented by a function:

7) $U: \{0,1\}^n \rightarrow \mathbb{R}$.

Values of different problems might be independent in the sense that U(s(t)) was always expressible as the sum of u-values for "elementary" solution vectors having only one non-0 component. We do not require this assumption. Intuitively, extrinsic values may be derived in some way from connections with other theory elements and/or technological considerations entirely extrinsic to "basic" T-science. We will return to consider their source below. Note that U(s(t)) represents the value of "having solutions to problems represented by s(t)"--not the value of some specific solutions. Note also that we consider the extrinsic values to be independent of the stage in the research program in which the problems are solved. U(s(tm)) is the value of having solutions to the problems s(tm) that terminates the research program s. Thus it is plausible to define the value of s to be

8) U(s) := U(s(tm)).

It might appear plausible to require further that the extrinsic value of having the solution to more problems is greater than that of having the solution to fewer. In view of 6), this would mean simply that, $U(s(t+1)) \geq U(s(t))$. We avoid taking this explicitly Panglossian view of the value of basic science.

Turning to probabilities, suppose that the probability of solution sequence s, given resource allocation vector r is given by:

9) $P(s,r) = P(\ s(t) \mid s(t-1),\ r(t)\ ) \times$

$$P(\ s(t-1) \mid s(t-2),\ r(t-1)\ ) \times$$

$$\vdots$$

$$P(\ s(1) \mid s(0),\ r(1).$$

That is, probabilities of problem solution histories are "Markov" in the sense that the probability of s(t) depends only on s(t-1) and the resource allocation vector at t. This is clearly an oversimplification in that it ignores investment in equipment at period $t_i$ that may be used for solving problems in periods $t \geq t_i$.

Finally, suppose that, the probability of problem solutions s(t), given the solution state s(t-1) and the resource allocation vector r(t):

10) $P(\ s(t) \mid s(t-1),\ r(t)\ )$

is a monotone non-decreasing, convex function of the resource allocation vector r, and nothing else. That is,

11) $P : S \times S \times R\ \ [0,1]$

is a conditional probability measure on S such that:

12a) if $s(t) \leq s(t-1)$, then, for all r(t),

$P(\ s(t) \mid s(t-1),\ r(t)\ ) = 0;$

12b) if $s(t) > s(t-1)$ and $r(t) > r'(t)$ then

$P(s(t) \mid s(t-1),\ r(t)\ ) \geq P(s(t) \mid s(t-1),\ r'(t)).$

That is, no allocation of resources can result in solved problems becoming unsolved (cf. condition 6) above) and more is always more effective than less.

Under these assumptions, the expected value E or a resource allocation vector r is

$$13) \quad E(r) = \left( \sum_{s \in S} U(s)P(s,r) \right) - L$$

and we might plausibly take this expectation value to be the objective function for an elementary research program in the environment of exogenously supplied values U(s). Recall that U(s) is the value of the problems solved at the termination of the program r. To conceptualize "intrinsic scientific value"--termed just 'scientific value' in what follows, we might simply suppose that the exogenous values are uniform--i.e. U(s) = U(S') = K, for all s S. Then

$$14) \quad E(r) = K \left( \sum_{s \in S} P(s,r) \right) - L.$$

Intuitively, we might think of K as being the value that some "disinterested," external benefactor attaches to T-basic science--"disinterested" in the sense that he/she does not care which T-problems get solved.

Assuming that r* is a research program that maximizes E(r) in 14), we might then identify r*(i,t) as the scientific value of having a solution to T-problem i at stage t. Intuitively, the scientific value of a problem in an elementary research program is just whatever resources would be allocated to its solution in an optimal research program. The content of this "definition" derives form the fact that we have a way of determining optimal research programs that is independent of knowing the scientific value of problems. On this account, the intrinsic scientific value of the solution to a problem is time dependent. I shall consider the intuitive plausibility of this below. In the case that there is more than one research program that maximizes E(r), r*(i,t) might not be uniquely determined in this way. At this stage of the analysis, I simply note this possibility and proceed to ignore it.

It is convenient to extend this idea to define the T-relative, or local, scientific value of temporal stages in solution sequences. The obvious way to do this is simply to sum be value of r*(i,t) over the problems with non-zero values in the vector s(t). Thus, I define

$$15) \quad VT_1(s(t)) := s(t) \bullet r^*(t) ,$$

where '•' is the scaler product.

The optimization problem posed by 14) can be attacked by the usual microeconomic methods. However, an intuitive understanding is perhaps more important for our purposes. The basic intuitive idea in 14) is that not all solution sequences are equally likely. There is generally a "natural order" in the solution of problems. Solutions to some problems become easier (i.e. cheaper), once other problems have been solved. Optimal research programs are going to be those which address problems in this natural order.

This is not a particularly novel nor profound idea, but, it does suggest a sharp, austere formulation of the concept of "scientific value" relative to a theory element T. Roughly, scientific value of unsolved problems is identified with the resources that would be allocated to their solution in some optimal research program in which external values are uniform. Intuitively, problems differ in scientific value because they differ in the degree to which they contribute to the solution to other problems. Still more intuitively, problems have scientific value just because their solution contributes to the solution of other problems. It is important to see that this concept of "relative" scientific value is independent of any assumption about the inherent social value of knowledge. Formally, this just means that it is independent of the value of K in 14). This independence is emphasized when we come to consider below how the probabilities in 14) might be determined solely by the formal properties of the theory element T.

With this intuitive understanding of our conception of scientific merit, it is not surprising that it turns out to be time dependent. When solutions to problems have only instrumental value as "means" to the solution of still other problems, it is natural to expect that this value will depend on which problems remain to be solved and, more specifically, which are next on the agenda of an optimal research program. One should expect that the scientific value of a problem is relative to the context of a specific research program <u>and</u> to a specific stage in that program. This suggests that optimizations in research planning should reassign priorities and reallocate resources frequently in the light of new information about the <u>actual</u> progress of problem solution. A more subtle analysis would countenance costs of information about the solution vector s as well as "reprogramming" costs.

This analysis of "scientific value" is relativized to a single theory element T simply because it is primarily with respect to single theory elements that our representation of scientific knowledge provides us with a characterization of scientific problems. It is also with respect to single theory elements that we have some clear idea of what might determine the probabilities of sequences of problem solutions. The most obvious idea is that these probabilities are determined by the constraints C that operate across sets of potential models. Intuitively, given some initial set of solved problems A, the most probable solution sequences are going to be those that add new members to A that are linked members of $A_t$ by constraints. For example, consider a research program aimed at applying Daltonian stoichiometry to a class of chemical reactions larger than that to which it has already been successfully applied. (For details on this example see, [1].) Common sense, historical data and formal analysis converge to suggest that the most economical research program begins with new reactions that involve at least some of the same compounds participating in the reactions to which the theory has already been successfully applied.

It is not obvious that constraints are the only factors relevant to probabilities of solution sequences for a single theory element T = <K,I>. The interpreting links that determine the intended applications I may operate in such a way that probabilities (as a function of cost) of acquiring data needed for problem solution differ among intended applications that are similarity situated with respect to the totally formal properties of T-constraints. For example, reliably pure samples of some, otherwise equally "interesting", compounds may differ widely in their cost. This suggests that our conception of an intertheoretical links might profitably be augmented with some kind of cost function.

Neither is it obvious how information about constraints, costs of interpreting links and perhaps other factors can be translated into probabilities of solution sequences. At best, these remarks suggest a way of posing the question.

## V. Basic Research Programs

Let us now consider how the concept of scientific value relative to an elementary research program for a theory element T may be extended to provide a more general concept of scientific value.

We may begin by generalizing our notion of an elementary research program. To do this, consider a situation in which a theory element T = <K,I> resides in an net of theory elements connected by interpreting links.

Suppose that, in the course of carrying out a T-elementary research program, additional data is required about member x of I. We represent the procedure of acquiring this data in the following way. Data about x is provided by potential model x'--linked to x--in some theory element T' that interprets T. Potential model x' may be among the solved problems of T'--$A'_t$. If so, the required data is just "exported" to T. If not, the x' is put on the "local research agenda" for T'. More precisely, the T'-scientific value of the problem x' has been augmented by virtue of its being called for as a part of the T local research program. Solving problem x' may require solving problems in other theories besides T that T' interprets and, as well, solving problems in theories that interpret T'. In this way, local research programs in T may be viewed as recursively generating local research programs in other theories linked with T. That is, global research programs can be viewed as resulting from aggregation of recursive propagation of local research programs. These ideas are made somewhat more precise in [9].

Given this picture of global research programs, an obviously attractive way to generalize our notion of scientific value is to envision that it propagates in a recursive and cumulative way through the net of theory elements. That is, the T-scientific value of the members of I in T = <K,I>--call it U[T] gets added to the T'-value of the members of I' in T' = <K',I'> when T' interprets T. Intuitively, members of I' are valuable as parts of elementary research programs relative to I' and valuable as well because they are linked with elementary research programs relative to T. In this way, members of I' would accumulate scientific value from all the theories that T' interpreted and, in turn, pass the <u>sum</u> of these values on to the intended applications in the theory elements that interpret T'.

Let us now consider the global scientific value--$V_g$--of a solution vector for T. The intuitive ideas just outlined suggest that $V_g$ should satisfy the following condition:

18) $$VT_g(s(t)) = VT_1(s(t)) + \sum_{link(s(t),s'(t))} VT'_g(s'(t)).$$

The notation 'link(s(t),s'(t))' means that s'(t) is some solution vector for a theory T' that T interprets and all non-zero components of s'(t) are linked to non-zero components of s(t).

Note that it is the global scientific value of the interpreted theories T' that appears on the left side of 18). Intuitively, this means that global scientific merit for T is accumulated from all the theory elements to which T is "forward connected" by interpreting links. This view has the intuitive advantage of making scientific value of a given problem a function of the "density" of the links between this problem and other problems (See [10].).

Roughly, this suggests that one begins calculating global scientific merit with those theories that do not interpret any other theories (if such exist). For these theories global and local scientific merit are identical. One then goes one level lower to the theories that interpret these and solves a maximization problem isomorphic to 13) with U(s) replaced by the $VT_g(s)$ given by 18). Then one iterates the same procedure at the next lower level until reaching the theory element of interest.

Aside from depending on heroic assumptions about our knowledge of the global structure of science and the implications of local structure (e.g., constraints) for solution probabilities, this account of global scientific value suffers from an additional defect. It depends on that assumption that interpretation links do not form closed loops in that procedures for evaluating the sum in 18) will only terminate under this assumption. It is far from obvious that this is true (See [1], Ch. 8). It would be nice to have some kind of measure of global scientific value that did not depend on this assumption. Intuitively, what's wrong with 18) is that theory elements "far away" from T contribute just as much as those "close to" T to the scientific value of problems in T. Adding some kind of "attenuating factor" like 1/(number of links in chain connecting T and T') might be more plausible. Still another approach would be to place some necessary conditions on global scientific value that related it to local scientific value via something like difference equations without necessarily determining its value. Intuitive considerations make this attractive, but do not appear to indicate an obvious way to proceed.

## VI. Special Laws and Law Discovery in Basic Research Programs

Up to this point, our discussion has focused exclusively on the solution of problems in an environment in which a single law is employed. Most scientific practice is not so simple. Theories with conceptual sophistication and complexity--for example classical equilibrium thermodynamics--typically consist of several theory elements each representing different laws linked by a "specialization links" into a tree like graph which coexists with the net of interpretation links. Specialization linked theory elements constitute a "theory element family" in that they all have the same conceptual apparatus $M_p$ non-theoretical conceptual apparatus $M_{pp}$. They differ only in their laws M and constraints C.

This specialization structure plays a role in the problems solving process in the single theory elements. Roughly, specialization provides additional possibilities for determining the values of theoretical concepts. These values may be passed through the specialization from one theory element net to facilitate the solution of problems in other theory elements. For example, masses discovered in Hooke's law systems (spring balances) may be employed in other Newtonian systems characterized by other force laws.

The account of basic research programs offered above needs to be emended to include this kind of activity. But, more importantly, this possibility suggests that a more radically different kind of scientific activity needs to be considered--the activity of discovering new theory elements related to existing ones by specialization links. On the account of "scientific problems" suggested above, discovering new scientific laws couched in a given conceptual framework is just a different kind of activity than discovering how to apply a known law to a new situation.

How can we appraise the scientific value of this "law searching" kind of activity? One approach it to regard it as "overhead" on the kind of problem solving activity we have already considered. The clue here is that having special laws always (in principle) or usually (in practice) makes the solution of these "first order" problems easier (cheaper). Formally, this results from their being additional possibilities of determining the values of theoretical concepts and transferring this information to where it can be used in (first order) problem solution.

Can this insight be translated, even in principle, into a quantitative measure of the value of having a new special law? For any specific additional special law, it is not implausible to suppose that one might quantify its contribution to (first order) problem solving. At least, this would be no more difficult than for constraints. However, the search for special laws is intuitively not an attempt to <u>justify</u> the addition of a specific, already formulated law. Rather, it is the search for <u>some</u> new special laws of unspecified form. How can we calculate the scientific value of such a search?

## Bibliography


[1] Balzer, W., C.-U. Moulines, and J. D. Sneed: 1987, An Architectonic for Science. Reidel.

[2] Gieryn, Thomas F.: 1978, "Problem Retention and Problem Change in Science," Sociological Inquiry, 48, 96-115.

[3] Hobbs J. R., and R. C. Moore: 1985, Formal Theories of the Common Sense World. Ablex, Norwood NJ.

[4] Krantz, D. L.: 1965, "Research Activity in 'Normal' and 'Anomalous' Areas," Journal of the Behavioral Sciences, 1, 39-42,.

[5] Kuhn, Thomas, H.: 1962, The Structure of Scientific Revolutions, University of Chicago Press, Chicago.

[6] Martin B. R. and J. Irvine: 1981, "Internal Criteria for Scientific Choice: An Evaluation of Research in High-Energy Physics Using Electron Accelerators," Minerva, 19, no. 3, 408-32.

[7] Merton, R. K.: 1957, "Priorities in Scientific Discovery: A Chapter in the Sociology of Science," American Sociological Review, 22, no. 6, 635-59.

[8] Mulkay, G. W.: 1975, "Problem Areas and Research Networks in Science," Journal of British Sociological Association.

[9] Sneed, J. D.: 1986, "Machine Models of the Growth of Knowledge: Theory Nets in PROLOG," to appear in Criticism and the Growth of Knowledge: 20 Years After: proceedings of an International Conference in Memory of Imre Lakatos, Thessalonokii, Greece.

[10] Weinberg, Alvin M.: 1984, "Values in Science: Unity as a Criterion of Scientific Choice," Minerva, 22, no. 1, 1-12.

[11] Weinstein, Deena: 1976, "Determinants of Problem Choice in Scientific Research," Sociological Symposium, 16, 13-23.

[12] Zuckerman, H.: 1978, "Theory choice and Problem Choice in Science," Sociological Inquiry, 48, 65-95.