Model-Based Evaluation of Long-Range Resource Allocation Plans

Ben P. Wise and Donald W. Kosy

CMU-RI-TR-85-22 (3)

Intelligent Systems Laboratory
The Robotics Institute
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

December 1985

To appear in *Artificial Intelligence in Economics and Management,* L. F. Pau (ed.), North-Holland P Co. (in press). Support for this research was provided by The Robotics Institute and the Eastern E Systems Company.

Table of Contents

Abstract

ıen corporate planning is decentralized, the plans produced by each suborganization must /iewed and evaluated to make sure they are reasonable and acceptable to the organization a ıole. In this paper we consider three ways of automating the evaluation task: two based on ru mbined with qualitative arithmetic, and one based on a microeconomic model combined w antitative reasoning and a search procedure. We argue that the knowledge encoded in the ru n be represented better using the model and that the search strategy implicit in the r presentation can be duplicated by the procedure. Moreover, quantitative reasoning can deal w nforcing and counteracting effects, while qualitative arithmetic cannot. This approach has be ed as the basis for the REMUS module in the ROME system.

ubtasks corresponding to the different levels of the organizational hierarchy. As estimates at low evels are generated, they are reviewed and then consolidated into estimates for the parent level. pproved at the parent level, these estimates become plans which guide future resource allocati ecisions.

While considerable attention has been given to methods for generating resource plans (f xample, mathematical programming), not much has been given to methods for reviewing pproving them. Traditionally, these activities have been performed by planning managers who a eld accountable for the consistency, completeness, and overall acceptability of plans made for the rganizational units. We were asked to develop a knowledge-based system to assist in the plan revie ctivity; this paper describes what we learned from that project and what resulted from it. The mo teresting result was that it was possible to recast the knowledge initially expressed to us in a lar umber of if-then rules into a general procedure plus a declarative model based on econom roduction functions. Using this knowledge in conjunction with a fairly general evaluation strate as sufficient to reproduce most of the expert behavior we observed.

. EVALUATING RESOURCE PLANS

To develop an appropriate algorithm and knowledge representation, we began with verbal protoco icited from a manufacturing plant planning manager as he reviewed several resource plans. Th ans were displayed as spreadsheets of numbers where the columns specified planning perio iscal years) and the rows showed the projections for each type of resource for each period. The ere also rows containing projections of output levels, inventory levels, breakdowns of shipments pe of customer, etc. (see [5] for details). During the course of the reviews, the manager describ hat features of the plans caught his eye, which seemed odd, what he had expected to find, how xplained away some oddities, and so forth.

His overall goal in this activity was to evaluate a plan against three basic criteria. The first w redibility: projected requirements should be consistent with known relationships and limits. F xample, an increase in output with no corresponding increase in raw material purchase impli omething strange is going on. The second was responsiveness: projected performance measur hould meet goals. E.g., a shortfall in desired productivity level means that either the goal nattainable or someone did not pay enough attention to it. The third was completeness: all resour eeds should be properly included in the plan. For example, an increase in projected productiv ithout some projected expenditure on process improvement almost certainly implies somebo rgot to include it in the plan. In short, every effect should have a cause, the effects should be tho esired, and all effects of a particular cause should be accounted for.[1] For each item in the plan, th anager's evaluation proceeded as follows. First, if there were any corporate goals for the item, ould check the values shown against the goals, since this was the easiest criteria to apply. If a go as not met, he would note it and continue. Often, however, there would be no specific goal for em. He would then check whether the values seemed "reasonable" relative to values of other item

[1] In addition, if the basic criteria were met, the manager was also interested in what the projections implied about resour quisition or disposal. We will not be concerned here with this aspect of the review.

To do that, he would first characterize the given item by various properties, such as whether t alues seemed high or low, or whether the trend was increasing or decreasing. From this, he wou ypothesize properties that other items should have to be consistent with this one. Finally, he wou heck the other items to see whether they in fact had such properties.

If all expected properties were found, then the manager would conclude that the original item w robably alright and move on. If one were absent, however, he would then begin checking t ssumptions implicit in his initial hypothesis. These assumptions typically involved items not shown ie plan, such as proportionality coefficients and minor terms in an aggregate sum. The usu ssumption was that these had remained constant. For example, if *Output* rose, then he wou xpect the number of *production-employees* to also rise, assuming that the two would maintain onstant ratio. If that expectation failed, he would then check the assumption by looking for xpenditure on process improvement, which would account for how the output per person ratio cou ave increased. If, in turn, that failed, the original item was taken to be problematic and was plac n a list of concerns to be taken up with the planner responsible.

The links between properties were originally classified as two separate rules of the form *if X, then* nd *if X and not-Z, then Y*. In practice, *X*, *Y*, and *Z* were conjunctions, as in the rule *if productio mployees rose and output rose, then expect material-spending up and inventory up*. (He expect iventory to rise because of a policy that inventory was to be held at approximately one fifth utput.) Noting this pattern of pairs of rules, we were led to adopt Doyle's formalism of *if X, unless ien Z* rules, rather than the usual *if X, then Z* rules [4]. Sixty-three such rules were found, enough eproduce the observed behavior. Doyle's formalism differs from *if (X and not-Y), then Z* in that t tter must explicitly, and perhaps slowly, verify that Y is false. The former checks only whether Y h ready been proven false. Only if Z turns out to be false will a thorough check of Y's truth or falsity ndertaken. If Y is chosen carefully, this not only allows non-monotonic reasoning, as Doyle explair ut can also be faster.

This rule-set had three severe—and related—problems:

- it was far too small to cover the space of expectations that are possible;
- the knowledge expressed did not seem to be the "basic" knowledge involved;
- its reliance on qualitative, rather than quantitative, description rendered it quite weak.

rst, in a rule like *if X, unless Y, then Z*, each variable in *X*, *Y*, and *Z* could be either positive, zero, egative (+,0,-), and either rising, constant, or falling (r,c,f). Clearly, the number of possib ombinations of signs and slopes, each with its own rule, climbs exponentially. Although our set xty-three was enough to reproduce observed behavior, it was only a tiny fraction of those that mig e needed to evaluate plans beyond the ones considered in the protocols.

Second, the rules did not seem to represent empirical correlations, as in a system like MYCIN [3], tuation-action pairs, as in R1 [9]. Rather, every rule seemed to have a "deeper" justification rms of known relationships among manufacturing resource variables. For instance, the justificatic se, then A will also rise. In this case, qualitative arithmetic (to be abbreviated henceforth as QA) nambiguous: (r) + (r) = (r). But if B rises and C falls, then A could do anything: (r) + (f) = (?), where neans 'unknown'. Thus, any extended chain of reasoning with QA may rapidly be dominated by (?) lowever, our planning manager could easily see that if B rose by 20, and C fell by 1, then A shou se, and he would rely on this inference subsequently. Hence, we began working directly w umbers and quantitative equations.

The use of equations solves the three problems just mentioned. First, since the relationships a ully spelled out, almost all the permutations of the qualitative rules can be automatically subsumed rdinary arithmetic. Second, since equations are essentially statements of facts about the activiti epresented by the variables in them, they constitute the "deeper" knowledge that motivated t riginal rules. Third, the effects of differences in magnitudes can readily be propagated along a cha f equations, again, by arithmetic. However, using equations introduces three new problems relat knowledge acquisition: (1) what functional forms do these equations have? (2) what are th arameter values? and (3) aren't equations too precise to reproduce the kind of reasoning peop se? Our solutions to these problems are presented in the next two sections.

. A MANUFACTURING RESOURCE UTILIZATION MODEL

Our first step was to divide the activities in a manufacturing plant into three generic subactivitie nanufacture of commodities, acquisition of new capital, and technical improvements to existi apital, as in the figure below. These subactivities were each represented as linear relations betwe uantities of inputs and of output, i.e. as Leontief production functions [8]. While fairly simple, the eontief functions were sufficient to reproduce our planning manager's reasoning.

The most important process is the manufacture of final product. The elements of the input vect Q, are machine-hours, floor-space-hours, materials-spending and man-hours; they are linea elated to output, Q, subject to capacity constraint, C, by the relations of equation 1, where α is t ector of resource-per-unit requirements.

$$I_Q = \alpha * Q, \; Q < C$$

Growth and technical change are represented as two separate processes, outside the producti rocess, for two reasons. First, they are not necessary for output; a factory can produce produc ithout process improvements, but not without raw materials. Second, they produce not mater utputs but changes in the parameters of the production process. Technical change alters t -vector; growth alters the capacity constraint.

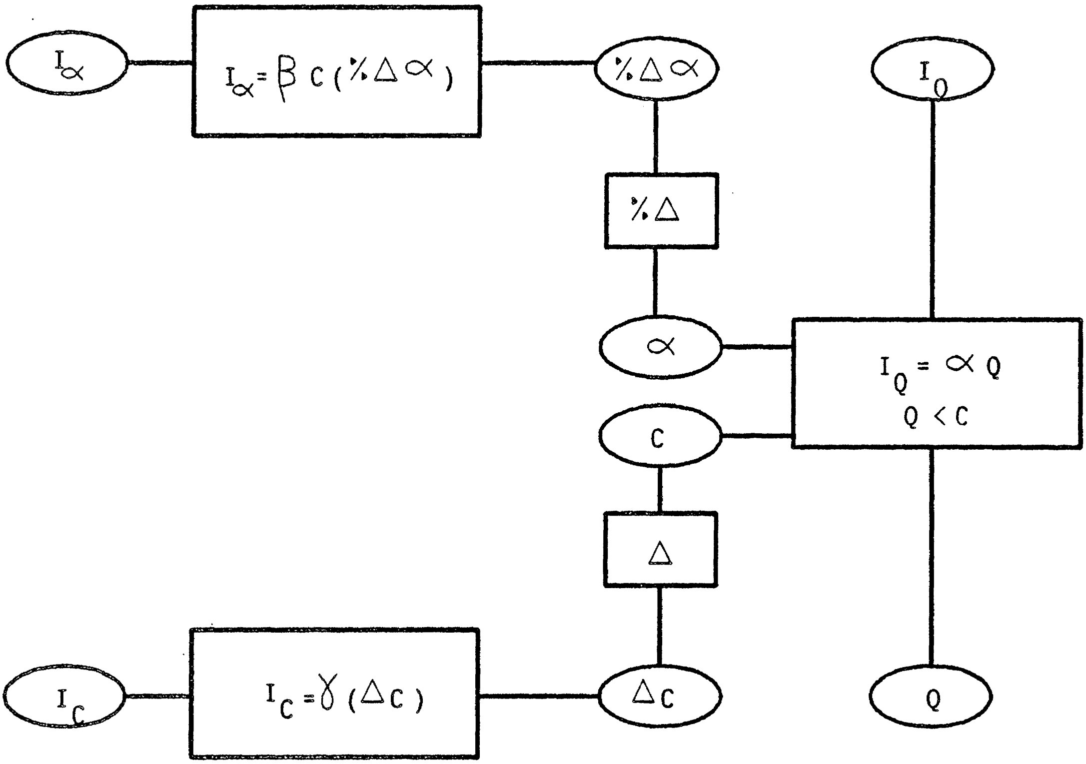
Below is the equation for technical change. It is linear in the percent improvement, and in t mount of equipment improved:

$$I_\alpha = \beta * (C * \Delta\alpha_i/\alpha_i)$$

the equation for capital growth, ΔC is the change in capacity, and I_C is a vector with elements direct labor for design and installation, and direct capital costs:

$$I_C = \gamma * (\Delta C)$$

quations were also added for cost-accounting, measuring performance, and miscellaneous relatio



The manufacturing, growth, and change models

‹e volume discounting of purchases, increased pay for overtime, the learning-curve effects ‹ roductivity, and so on. In effect, these were a financial scaffolding built on top of a three-pa hysical model.

The above three equations bring up a difficult, fundamental question: since the parameter valu id not explicitly appear in the plans, how could we get numerical values for the parameter vectors , γ? Our answer comes from the observation that these values embody expertise in much the san ense that rule-strengths embody expertise in rule-based expert systems. This analogy suggests th ey might be elicited directly. Or, they might be estimated statistically from data, such as past histo r the current plan. Estimation from historical data might reveal, for example, that the current pl ssumes unusual productivity over its entire span. Estimation from the current plan might reve utliers.

In our case, however, the estimation process was trivial. Taking a cue from our human reviewer, v mply used values of parameters taken from current plant operations, which were readily availab f course these would be quite unsuitable if we tried to use them to generate resource projectior

Given the set of equations and parameter values just described, the linkage between propert
nitially represented by rules can now be made by arithmetic operations on variables. In particular
now possible to determine the influence of variables on each other by tracing through the graph
elationships induced by the equations. However, since the human reviewer was only concerned w
he most significant influences, it is important to be able to disregard "insignificant" effects during t
easoning process. To do that, we have developed a numeric influence measure called ε which te
ow much a set of variables effects a given variable, in the change between two contexts.
= F(B,C), and the two contexts are 1 and 2, with $A_1 = F(B_1,C_1)$ and $A_2 = F(B_2,C_2)$, then B's effect on
defined by equation 4.

$$\varepsilon(A,\{B\}) = F(B_2,C_2) - F(B_1,C_2)$$

ikewise, the effect of both variables, {B C}, is defined by equation 5:

$$\varepsilon(A,\{B\ C\}) = F(B_2,C_2) - F(B_1,C_1)$$

he two contexts could be two different years of a plan, or actual versus goal, or any other pair
ontexts for which the variables were all related by the same F. To drop insignificant effects, and al
low for competing effects, we did not require a perfect match between ε and the the change in
A. Rather, we said that a set X accounted for ΔA if inequality (6) held, where τ was empirically set
8.

$$1/\tau > \varepsilon(A,X)/\Delta A > \tau$$

nlike correlation coefficients, this measure is still accurate for nonlinear functional relationshi
nlike partial derivatives, it does not presume that other variables remain constant in determining t
fluence of a given variable. The slack in (6) also allows minor estimation inaccuracies to pa
nnoticed. Further details may be found in Kosy & Wise [6].

A procedure for using this measure that matches the planning manager's review strategy can
ated as follows. Given a property of one variable, start tracing back, breadth first, via all equatio
volving that variable, collecting the influences that account for the property, until an "adequa
ause" is found. If the tracing requires checking a variable whose value is not shown on the pla
nd a suitable surrogate, and check it, where a surrogate is a variable whose value is shown that
nctionally related to the missing variable. If a cause is found, but the relevant variable is not shov
the plan, check to make sure that its other expected effects, beside the initial property, a
esent.

One additional kind of knowledge must be included in a model in order to use this procedure to fi
"adequate cause" of observed properties. A list of equations is not inherently directional, b
ausality is. Hence it is necessary to provide an ordering on the variables that indicates th
recedence. This precedence was implemented by labeling equations as 'aggregation', 'definitio
olicy', or 'physical constraint'. An example aggregation would be the relation between to
roductivity and the various factor productivities, or between total labor and each sub-type of labo
n example definition would be that total productivity is the amount of output over the total value of
puts. An example policy would be that inventory should be 20% of the yearly output, or th
roductivity should improve according to a learning curve with a pre-planned rate of improveme
ach of the three basic manufacturing plant processes are 'physical constraints'. The 'aggregatio
nd 'definition' labels serve as flags to make sure that the superordinate variable is explained in terr
the subordinate variables, not vice versa. If a variable is connected to a 'policy' equation, there

lentify the end of a causal chain, certain variables were labeled as exogenous. No attempt is ma) find causes for properties of these variables since they are taken to be outside the system bei nodeled.

As a very simple illustration of how this procedure reproduces the planning manager's reasonin onsider the rule *if employee-hours up, then expect material-spending up and output up*. On employee-hours up" has been noticed, that is traced to the exogenous variable, Q. However, ariable in the plan shows the value of Q in terms of volume of output produced. Rather, what nown is the dollar volume of output, which is related to Q by price. Hence, the procedure checks t ollar output variable and, as a double-check, the other inputs (e.g., material-spending) required roduce plant output are also checked.

A more complex example is *if total-space up, employee-hours constant, and distribution-spa onstant, then expect production-space up, output up, and capital-for-operations-improveme ositive*. First, since total-space is defined as the sum of production-space and distribution-spac ıen the total rising with distribution-space constant requires production-space to rise. Second, wh ne input factor (space) rose, others (employee-hours) have stayed constant, hence the α-vect ıust have changed, and so the trace would proceed to examine inputs to the technical chan rocess. In the original rule-set, the presence or absence of countering effects had been reflected sing multiple rules, one for each combination of missing and countering effects. These were lat ıcorporated into the "unless.." clauses. For example, if *output* rose, increases in each input we xpected. If output rose and an input did not rise, then a change in the α-vector was sought, becau ıat is the only remaining term in $\alpha_i{}^*Q = I_i$. Because α's were not listed in the plan, expenditure f apital improvements was checked, because that was a reliable surrogate, and hence a good che ır whether the expected countering effect was actually present. Thus, the elicited rule actua mbodies two distinct analyses, which the above procedure reproduces in sequence.

Clearly the graph-searching strategy can trace whatever combinations are present, without needi special rule for each one. However, one should note that this procedure is not exactly the san ing as what the manager described. That is, when asked how he reviewed resource plans, he ga any specific rules. Our general procedure traces influences, and mentions in sequence tho ariables that appear on the plan as it encounters them in the equational model. These turn out to e same variables that the planning manager mentions in his rules.

Our quantitative approach to explaining numerical results contrasts strongly, and instructively, wi ıat of Bouwman [1,2]. Bouwman's domain was also quantitative data about a manufacturi oncern, but the reasoning was performed entirely in QA. His system basically scans the data, an ır each equation in the model, forms an expectation about the term it defines. For example, $= B + C$, B rose, and C rose, then A is expected to rise also. These expectations were then fitt gether into chains, and those confirmed by the data were cited as explanations, discarding tho ontradicted by the data or simply not addressed by it. Our approach differs in that it begins at ariable and works out from there, rather than doing a complete scan before putting the piec gether. More importantly, Bouwman did not include cases where QA gives ambiguous resul /hile this strategy does prevent domination by (?)'s, it also rules out more accurate analyses whe

and C, unless D were omitted from the equation entirely. But then, it would be misled in those cas
here D actually was important. Thus, Bouwman's approach is more dependent on a powerful a arsimonious model of a firm than is ours. It relies on the significant factors to have already be icked out when the model was specified, as it has no means of picking them out on a case-by-ca asis. In the resource plans we examined, however, the relative magnitudes of the variables we rucial in deciding which lower level variables to explore next.

In summary, we discovered that the original sixty-odd highly specific rules could be generated eneral knowledge about arithmetic applied repeatedly, in different ways, to a more or less gene iodel of how resource variables are related in a manufacturing plant, as expressed by simp quations. This is in marked contrast to a system like MYCIN, where rules express empiric orrelation but are not linked to a model. The use of arithmetic contrasts with a system li ouwman's, which reasons purely qualitatively, not quantitatively.

. THE REMUS EVALUATION PROCEDURE

A version of the model and reasoning procedure just described has been adapted for use in t EMUS module of the ROME system[7]. ROME is an experimental decision support system generat iat allows specification of planning models in terms of variables and algebraic formulas. Models m e executed and the results displayed in spreadsheet form. The purpose of REMUS is to Review a valuate a Model's Underlying Structure by comparing results to evaluation criteria. Gene iodels—such as the one for manufacturing resource utilization—and quantitative reasoning are us support the evaluation process.

Two sorts of evaluation criteria may be declared: norms and goals. A *norm* is a relationship amo ariables that "should" hold true under "normal" circumstances according to experts in the doma f the model. A *goal* is a statment of an organizational objective or policy that is likewise expressib terms of model variables. Norms and goals are declared in ROME by *expect* and *want* statemen espectively.

To illustrate, the three production functions in the manufacturing model are relationships th iould hold true in a manufacturing plant. Expanding out the vectors in equation 1, and usi omewhat more mnemonic variable names, yields the following set of norms:

```
Expect Direct Labor Hours to equal Labor Hours/Unit * Q-produced
Expect Total Materials Cost to equal Material $/Unit * Q-produced
Expect Total Floorspace Hours to equal Floorspace Hours/Unit * Q-produced
Expect Total Machine Hours to equal Machine Hours/Unit * Q-produced
Expect Q-produced to be no more than Plant Capacity
```

quation 2 has two components which may be expressed as follows:

```
Expect abs(%change(Labor Hours/Unit(y))) to equal Project Hours(y-1) /
        Plant Capacity(y) * Labor Hours per Unit Productivity Change(y-1)
Expect %change(Labor Hours/Unit(y)) to equal Addon Capital Expense (y-1)/
        Plant Capacity(y) * Capital Cost per Unit Productivity Change(y-1)
```

quation 3 may be expressed similarly.

Regarding goals, there may of course be many for a particular plant at a particular time. Howeve roducts should be manutactured more efficiently than comparable current products, or at least i iss efficiently. This goal, applied to labor efficiency, may be expressed to ROME as:

```
Declare comparable product to be an anchor variable
Declare current year to be a column variable
We want Labor Hours/Unit to be no more than Labor Hours/Unit for a
        comparable product in the current year
```

he second goal is that productivity should improve at a desired rate as a function of the amount reduction experience that has accumulated for a given product This relationship is known as earning curve', and one way to state the goal is:

```
Define Cumulated-Q(y) to be Cumulated-Q(y-l) + Q-produced(y)
Define Cumulated Direct Labor Hours(y) to be Cumulated Direct Labor Hours(y-l)
        + Direct Labor Hours(y)
Declare yl to be a column variable
Define yl to be the first year of production of the product
Let Cumulated-Q(yl-l) ‑ 0
Let Cumulated Direct Labor Hours(yl-l) - 0
Define %Additiona1-Q(y) to be Cumulated-Q(y) / Q-produced(yl)
Define LCRF to be the desired learning curve cost reduction factor
Define Alpha to be log2(LCRF)
We want Direct Labor Hours(y) to be no more than
        Cumulated-Q(y) * "/Additional -Q(y) t Alpha * Labor Hours/Unit(yl)
        - Cumulated Direct Labor Hours(y-l)
```

D actually check these goals, their parameter values (e.g., comparable product, current year, LCR ust be set We will do this in a moment.

The main steps in the REMUS evaluation procedure that applies the above criteria are shown belo D evaluate a variable *v:*

1. If there are criteria for the values of v, apply them and state conclusions. Check equalities before inequalities.
2. Find variables that explain the trend in *v* using the formula for *v* given in the planning model. If there is no formula but there is an equality norm in the generic model, use that. Call this set of variables S.
3. If there are other variables in the formula which have evaluation criteria but are not in S, add them to S.
4. Recursively evaluate the variables in S.
5. If there is no formula for *v* (S is empty), go on the the next variable to be evaluated.

iis procedure may be continued until it reaches exogenous variables. The main conclusions th ay be drawn are as follows:

1. If criteria are met? say so.
2. If a goal is not met, conclude that *v* is 'problematic'.
3. If there is a moderate difference (< 30%) between a value and a norm, conclude that *v* is 'OK' with respect to that norm.
4. If there is a large (> 30%) difference between a value and a goal, it is 'extraordinary'. If there is a large deviation from a norm, it is 'odd'.
5. If the effect on *v* of a large deviation in a lower level variable can be determined, suggest that *v* may be 'too high' or 'too low'.

ie boundary between moderate and large differences is taken from Bouwman [1], This procedure differs from the one described in the previous section in several respects. Fir riteria are applied whenever possible. This tends to focus attention on every issue that may affe ie top level variable being evaluated. Second, all explanatory variables are found at once, using t measure and test 6. This contrasts with the hypothesize / check assumptions / find counteri ictor strategy. Moreover, explanation paths are pursued depth first, rather than breadth first. Bo f these make it easier to express each step of the evaluation in words. Finally, REMUS makes efault assumptions about the values of variables not shown on a plan. Rather, these assumptio iust be stated explicitly, if needed. The reason for this is to make sure the user knows wl ssumptions are being used, and to sidestep the problem of choosing good surrogate variables.

One thing remains to be done before we can use this procedure to evaluate a specific resource pl nd that is to link the criteria to the particular variables in that plan. Consider the portion of esource plan shown below:

redwood	1985	1986	1987	1988	1989	1990
production spending	1.40	4.87	4.79	3.45	4.73	5.85
mfg dl $	0.0	2.49	4.08	2.86	3.90	4.94
other dl $	1.37	2.38	.71	.58	.82	.92
mfg employee-hours	0.0	44.40	73.60	50.79	62.68	77.21
other employee-hours	31.40	42.50	12.80	10.30	13.20	14.40
labor rate	44.47	56.00	55.44	56.38	62.27	64.02
process improvement hrs	0.0	0.0	0.0	0.0	0.0	0.0
capital for improvement	0.0	0.0	0.0	0.0	0.0	0.0
output $M	0.0	36.75	62.16	52.06	65.36	81.13
avg price		.25	.21	.19	.19	.19

Redwood" is the name of a new product that is to be introduced in 1986. The relationships in t anning model that generated these results are as follows:

```
Define production spending to be mfg dl $ + other dl $
Estimate mfg dl $ to be mfg employee-hours * labor rate / 1000
Estimate other dl $ to be other employee-hours * labor rate / 1000
```

ll the other values are input. Dollar values are expressed in millions except for labor rate, and lab ours are expressed in thousands.

The variables in this plan may be linked to variables in the generic model given above by declari em to be "kinds" or "measures" of the corresponding generic variables. The norms and goals fro e generic model will then become norms and goals for the planning variables by a process nheritance" described in [7]. The ROME statements expressing these correspondences are:

```
Declare mfg employee-hours to be a kind of Direct Labor Hours
Declare capital for improvement to be a kind of Addon Capital Expense
Declare process improvement hrs to be a kind of Project Hours
Assume Q-produced to equal output $M / avg price
Declare labor rate, Labor Hours/Unit and Q-produced to be exogenous
Define Labor Hours/Unit to be Direct Labor Hours / Q-produced
```

everal other facts that bear on the evaluation are also known. The current product being produc 1985 is called "Hemlock" and the labor needed to manufacture it is 640 hours/unit. For the ne roduct, Redwood, being introduced in 1986, it is desired that productivity improve by 10% for ea

```
Let comparable product(redwood) = hemlock
Define yl(redwood) to be 1986
We want LCRF(redwood) to be .9
Let Labor Hours per Unit Productivity Change(redwood) =
        Labor Hours per Unit Productivity Change(hemlock)
Let Capital Cost per Unit Productivity Change(redwood) =
        Capital Cost per Unit Productivity Change{hemlock)
Let Plant Capacity(redwood) = Plant Capacity(hemlock)
```

We are now in a position to evaluate several variables in the plan. However, we will only show /aluation for one - production spending - since the others will be evaluated along the way:

```
> Evaluate production spending(redwood) for 1989
There are no evaluation criteria for production spending(redwood) in 1989, However,
redwood production spending goes up in 1989 because mfg dl S and other dl $ both go
up and production spending = mfg dl $ + other dl $.  The increase in production
spending was due primarily to the increase in mfg dl $. Would you like me to
continue? Y

Redwood mfg dl $ goes up in 1989 because mfg employee-hours and labor rate both go up
and mfg dl $ = mfg employee-hoars * labor rate / 1000.  The increase in mfg dl $ was
due primarily to the increase in mfg employee-hours.  Would you like me to continue? '

By norm 1_9 we expect redwood mfg employee-hours to equal Labor Hours/Unit * Q-producei
Assuming Q-produced ^s output $M / avg price, this appears to be so in 1989.

By goal 2, we want redwood mfg employee-hours to be no more than Cumulated-Q(y) *
7oAdditional-Q{y) t Alpha * Labor Hours/Unit(yl) - Cumulated Direct Labor Hours(y-l).
This appears to be so in 1989.  But we also find redwood mfg employee-hours to be
much less than that value in that year.  This seems extraordinary.  So redwood
mfg employee-hours may be too low in 1989, which means mfg dl $ may be too low, which
means production spending may be too low.  Would you like me to continue? Y

By norm 6, we expect abs(%change(redwood Labor Hours/Unit(y))) to equal process
improvement hrs(y-l) ⁄ Plant Capacity(y) * Labor Hours per Unit Productivity
Change(y-l).  This seems to be OK in 1989.

By norm 7, we expect %change(redwood Labor Hours/Unit(y)) to equal capital for
improvement(y-l) / Plant Capacity(y) * Capital Cost per Unit Productivity Change(y-l)
This seems to be OK in 1989.

By goal 1, we want redwood Labor Hours/Unit to be no more than Labor Hours/Unit
(hemlock,1985).  This appears to be so in 1989.  But we also find redwood Labor
Hours/Unit to be much less than that value in that year.  This seems extraordinary.
So, again, mfg employee-hours may be too low in 1989, which means mfg dl $ may be too
low, which means production spending may be too low.  Would you like me to continue?
```

ie procedure would now go on to evaluate Q-produced, labor rate, and other dl $. It can be se)m the above, however, that the 1989 projection for mfg employee-hours is quite optimistic. This rn, diminishes the credibility of the mfg dl $ and production spending values that depend on it

. CONCLUSIONS

Evaluating a resource plan may be viewed as a process by which a reviewer compares what ıows or believes to what he is told by the numbers in a plan. For a computer program to perfo ich comparisons, it must use some representation of a reviewer's knowledge and beliefs. While c itial representation of this knowledge was in terms of if-then rules, we discovered that relatio

Although, in this paper, we have confined our attention to evaluating resource plans, the issu volved are very much related to the general problem of model validation. Hence, it may not rprising that parts of the evaluation strategy we induced from a human reviewer's protocols nilar to more formal model validation techniques. For example, in hypothesizing variables t ould change when a change is observed in a given variable, a human reviewer is performing a k sensitivity test. In comparing the value of a variable against what that value would be for curr ant operations, he is essentially trying to verify that the planning model would produce corr sults for a known test case. Finally, the entire strategy rests on the principle that values can best dged by comparing them to other values that have independent justification.

The chief strength of the REMUS procedure is that it allows evaluation of a result even if there criteria directly applicable to it. By descending to subordinate variables that affect a giv riable, this procedure conducts a search for criteria that may be relevant based on functio lationships. Moreover, because the variables in a given planning model can be linked to expect lationships, a given set of criteria can be transferred to other models in the same domain. Togeth ese mechanisms provide a means of using relatively general, structural models to make judgme out specific cases.

EFERENCES

Bouwman, M.J., *Financial Diagnosis: A Cognitive Model of the Processes Involved*, PhD Thes Graduate School of Industrial Administration, Carnegie-Mellon University, 1978.

Bouwman, M.J., "Human Diagnostic Reasoning by Computer: An Illustration from Financ Analysis," *Management Science*, V. 29, No. 6, pp. 653-672, June 1983.

Davis, R., *Applications of Meta Level Knowledge to the Construction, Maintenance and Use Large Knowledge Bases*, PhD Thesis, Computer Science Department, Stanford University, J 1976.

Doyle, J., "Methodological Simplicity in Expert System Construction: Judgments and Reasor Assumptions," *The AI Magazine*, V. 4, No. 2, pp. 39-43, Summer 1983.

Kosy, D.W., and Dahr, V., "Knowledge-Based Support Systems for Long Range Planning," T Robotics Institute, Carnegie-Mellon University, 1983.

Kosy, D.W., and Wise, B.P., "Self-Explanatory Financial Planning Models," *Proceedings AAAI-84*, pp. 176-181, August 1984.

Kosy, D.W., and Wise, B.P., "Overview of ROME: A Reason-Oriented Modeling Environmer The Robotics Institute, Carnegie-Mellon University, 1985.

Leontief, W., et al., *Studies in the Structure of the American Economy*, New York: Oxfo University Press, 1954.

McDermott, J., "R1: A Rule-Based Configurer of Computer Systems," *Artifial Intelligence*, V. No. 1, pp. 39-88, September 1982.