10

INTEGRATIVE POPULATION ANALYSIS FOR BETTER SOLUTIONS TO LARGE-SCALE MATHEMATICAL PROGRAMS

Fred Glover,
John Mulvey,* Dawei Bai,* Michael T. Tapia**

The University of Colorado at Boulder
Boulder, Colorado 80309
USA

**Department of Civil Engineering and Operations Research*
Princeton University
Princeton, New Jersey 08542
USA

***Department of Management Science and Information Systems*
The University of Texas at Austin
Austin, Texas 78712
USA

ABSTRACT

Integrative Population Analysis unties the learning process called target analysis and a generalized form of sensitivity analysis to yield improved approaches for optimization, particularly where problems from a particular domain must be solved repeatedly. The resulting framework introduces an adaptive design for mapping problems to groups, as a basis for identifying processes that permit problems within a given group to be solved more effectively. We focus in this paper on processes embodied in parameter-based definitions of *regionality*, accompanied by decision rules that extract *representative solutions* from given regions in order to yield effective advanced starting points for our solution methods. Applied to several industrial areas, our approach generates regions and representations that give an order of magnitude improvement in the time required to solve new problems that enter the population and therefore makes the application of large scale optimization models practical in reality.

G. Yu (ed.), Industrial Applications of Combinatorial Optimization, 212–239.

1 INTRODUCTION

Many industrial applications of optimization involve a series of interrelated problems that represent goals associated with different individuals, companies or scenarios. In addition, these interrelated problems are not just "one shot" occurrences but must be addressed over time. Integrative Population Analysis provides a framework that enables successively encountered problems in such applications to be solved more effectively and efficiently. Our approach integrates a refinement of the learning process of target analysis with a generalization of the mathematical programming process of sensitivity analysis.

Integrative Population Analysis characteristically deals with problems that belong to a common class or that arise from a common underlying model. The context for these problems has an importance that is sufficiently encompassing, or sufficiently enduring, to require a significant number of problem instances to be solved. The goal of providing improved methods to operate in such settings is far from new. However, we provide a new design for achieving this goal that proves remarkably effective, based on combining and extending principles from artificial intelligence and mathematical optimization.

Our approach begins by creating special groupings according to a notion of *parametric proximity*. The properties that characterize membership in a particular group are subject to redefinition over time and the parameters that define these properties are a function of both problem data and problem structure.

We may illustrate some of the relevant concerns by reference to several industrial applications. The first application is on flexible manufacturing systems. Consider the situation where products are to be manufactured each day according to specific requirements of customers. The associated machine scheduling problem has to be solved again and again whenever a new product is ordered. These machine scheduling problems differ in the values of their parameters, but otherwise exhibit similar structures. The solutions to some of these problems may resemble each other in various ways, while the solutions to other problems may be unrelated. Another application is on marketing strategy design where the preference patterns of customers need to be uncovered. In this case, we need to look at the population of problems that maximize individual customers' utility functions. Again these problems share similar structures but differ in the specification of preference functions and the available financial resources of individual customers. Finally, we can apply Integrative Population Analysis to product customization. This allows a firm to identify new products that can be created to service different customer segments. Common customer characteristics can be determined so that the a firm's existing strengths can be used to develop products that will attract new customers.

To take appropriate advantage of shared features of problems and solutions, we are prompted to ask the following questions. On what basis can we usefully characterize the individual problems as similar or dissimilar? How can we uncover properties of solutions or solution processes for similar problems that will allow us to solve their problems more effectively? How can we assure a design that will be robust and respond to changing conditions? These issues will be discussed in this paper both in a general context and in the context of the several industrial applications just described. The implications of our approach extend beyond the immediate boundaries of such applications, and are relevant to determining policies for supporting undertakings.

To set the stage for presenting the tools and special features of Integrative Population Analysis, we first examine some of the areas that create an important need for its applications. We then present a description of the key properties of IPA in Section 2, and discuss how it can be applied to enhance the computational efficiency of practical optimization processes in Section 3. We conclude with an examination of future directions in the final section.

1.1 Applications for IPA

IPA can be applied to any problem that has many instances that have to be efficiently solved. These problems can occur in many industrial situations, especially in industrial marketing and design. One area that motivates the need for Integrative Population Analysis is product differentiation, especially in an industrial setting. By an industry we mean a group of firms which produce goods that are close substitutes in consumption. Consumers choose goods from different firms based on combinations of certain product features. If a particular combination of features is highly attractive to customers but unavailable, new enterprises that are capable of providing such features are encouraged to enter the market. For this reason, companies that have monopoly power customarily seek to offer a wide range of product features to block opportunities for new entrants that may otherwise compete with them. The major players in the soft drink industry, for example, provide numerous variations of their basic products to serve people with different tastes. At the same time, a small start-up company that successfully perceives and responds to an opportunity overlooked by the entrenched leaders can sometimes reap considerable rewards. Examples of this abound in markets ranging all the way from ice cream to computer software. It is clearly important for companies, large or small, to analyze the market in order to survive and prosper. Integrative Population Analysis increases the scope and effectiveness of such studies.

Product differentiation [15, 14] can be expressed along many dimensions. For example, Golden Delicious apples differ from Macintosh apples in their juiciness, tartness, sweetness, color, texture, and so forth. When apples on the market neglect to offer a feature that a significant number of people prefer, these individuals will switch to a new brand that more adequately meets their preferences, thus creating an opening for such a brand to emerge. Lancaster [19] was the first to analyze such a situation.

To illustrate, and to establish connections with concerns that will shortly be elaborated in greater detail, we consider goods with only two characteristics, such as apples with characteristics Sw and So (Sweetness and Sourness). The first quadrant in Figure 1 depicts the space of characteristics that represent various types of apples. The quantity of each characteristic is measured along each axis. For example, type C Apple offers three units of sourness for every two units of sweetness. The other rays represent other types of apples which have different So/Sw ratios. Here we have five types.

An efficiency frontier identifies the maximum amounts of characteristics the consumer can obtain with a given level of expenditure. To derive this frontier, we require some standard assumptions that are not entirely realistic in all situations. Nevertheless, they are sufficiently well approximated under many conditions to allow us to rely on them for our preliminary analysis.

The first step is to find out how many units of each variety the consumer would obtain in separate transactions by devoting his total apple budget to buying a single variety at a time. This is found by dividing the apple budget by the price per apple in each case. The result gives a number of units which can be measured as a distance along each of the rays of Figure 1. These are shown as the points a, b, c, d and e in Figure 1.

The efficient frontier then joins successive points that lie on the outer perimeter to produce the convex hull of these points. Therefore b cannot be connected with a and c because it lies strictly inside this outer perimeter. In fact, Type B apple's price is so high that it will not be bought since a combination of A and C will get more units of each characteristic at point w, which lies on the frontier. (The ability to create a "composite good" that corresponds to w is one of the basic assumptions of such analysis.)

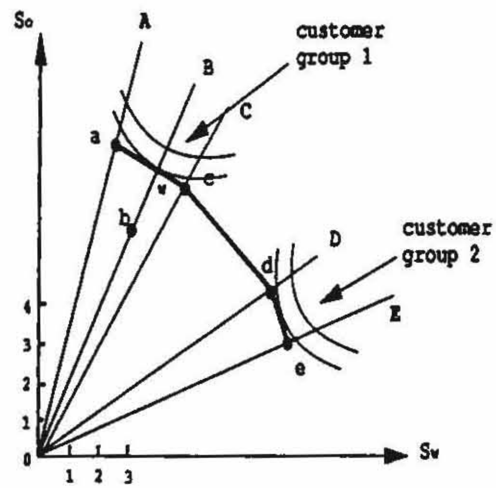

Figure 1 Utility maximization and market segmentation

Individuals possess preference structures which give rise to indifference curves as in Figure 1. Two sets of indifference curves are present, one for each customer group. The utility maximizing choice of consumption can be found by identifying the point at which the highest indifference curve is reached.

Individuals will usually have different preferences for different characteristics. Preferences for certain ratios of characteristics may cluster in one or more parts of the space, leaving other parts of the space largely unoccupied. When this occurs, we say the market *segments* (i.e., becomes segmented). Products outside these segments will find few customers. If the market does not exhibit such segmentation, that is, if customers are allocated uniformly along the efficient frontier, the frontier still provides useful information. We will show how to exploit these classical conditions, and others that are more general, in our subsequent development.

Continuing to lay the foundations that motivate the need for IPA, and that simultaneously introduce certain basic concepts to aid in understanding its characteristics, we turn next to the area of marketing strategy design. The first step is to understand the situation in which products will be accepted by customers, which requires an analysis of the nature of customer demands. This leads us to a characterization that divides the population into neighborhoods or cohorts.

For us, *cohorts* represent groups of customers who share certain common characteristics (see section 2.4). A cohort is constructed in a manner that enables us to evaluate the desirability of various products for *all* customers that lie within it. We may find, for example, that members of a particular cohort will greatly benefit by a new type of home mortgage, but will benefit much less from a traditional mortgage. Such relationships have the advantage of allowing us to design a marketing strategy that matches existing products with the cohort. More particularly, it becomes possible to creat optimization models to exploit market differentiations based on these relationships.

Another potential application area of IPA is in product customization, which is clearly related to product differentiation. For instance, different soft drinks are customized products that serve people with different tastes. An example that merits special consideration concerns investment. In the analysis of product differentiation, if the two attributes underlying the analysis illustrated in Figures 1-3 include elements such as mean and variance of a portfolio, we can then investigate patterns of demand for financial products with different returns and risks. The classical building blocks for generating and exploring efficient frontiers, which we link to customer preferences, can be embedded simply and advantageously in our processes for characterizing cohorts.

Once cohorts are defined via customer preferences, it becomes possible to "engineer" new products based on our analysis of the usefulness of the engineered product. In the investment area, for instance, there is great interest in financial engineering – the creation and trading of novel securities. Integrative Population Analysis provides a vehicle for assisting in the design of new products, by systematically evaluating the improvements for various segmented populations in ways not previously possible.

1.2 Computational Speed of IPA

Optimization problems in many practical settings are large or have attendant complexities such as nonlinearities, combinatorial structures and uncertainties. Size and complexity are often related, as where uncertainties can translate into special nonlinearities whose treatment includes the introduction of numerous variables and constraints. The solution of real world models in such settings typically involves a substantial investment of computation time. We begin by focusing on the goal of reducing the real-time computing burden for solving such problems.

To handle this situation with Integrative Population Analysis, we first face the challenge of mapping different problems into groups so that we can design a way to treat them more effectively based on this differentiation. In our present development, for concreteness, we will devote our attention to a mapping strategy that derives from identifying multiple elite solutions to the problems considered, and grouping problems both in terms of their data parameters and in terms of certain characteristics of these elite solutions. These elite solutions allow us to more efficiently solve similar large-scale mathematical programming problems by providing us with better warm-start solutions. We stress that these characteristics are not defined in advance, and are not simply *descriptive* but *functional.* Specifically, the determination of membership within a common group will depend on the existence of a *community of elite solutions*, derived as a subset of the union of their separate elite solutions, such that problems in the group can be solved more efficiently by utilizing members of this elite community. Note the composition of an elite community rests on several notions that must be clarified. We must specify, for example, what we mean by "utilizing" members of the elite community. Subject to such clarifications, we see that problems of a common group may in fact vary substantially on many descriptive dimensions. Our concern is to unite them at the functional level of identifying ways to exploit them more effectively. At the same time, we anticipate that a connection exists between various problem parameters and such functional concerns, and we will show how to integrate the functional and parametric considerations in our following development.

2 FUNDAMENTAL CHARACTERISTICS OF INTEGRATIVE POPULATION ANALYSIS

2.1 Population Environment

There are three environments for conducting our analysis. In the first situation, data for all the individuals of the population are known a priori. A thorough study can then be performed beforehand. In the second situation, data for the individuals arrive sequentially. We then must determine how to take appropriate account of a growing body of information. In the third situation, data for the individuals are unknown but sampling is possible. This leads to consideration of effective procedures for selecting such samples.

2.2 Problem Formulation

Let P be a population of agents (customers, investors, etc.). Each agent, $p \in P$, has the following optimization problem to solve:

$$\begin{array}{ll} \max & f_p(x_p) \\ \text{s.t.} & x_p \in X_p \end{array} \tag{10.1}$$

Let x_p^* be an optimal solution to the problem and let $f_p^* = f_p(x_p^*)$ be its optimum objective value. For example, $f(x_p^*)$ represents a consumer's preference function at an optimal solution. We assume that these optimization problems are structurally related, as by representing instances of a particular problem class with different data sets.

For a variety of applications, the feasible region X_p in (10.1) can be defined as a collection of linear inequalities summarized by:

$$A_p x_p \;\leq\; b_p \tag{10.2}$$

where A_p is an $m \times n$ matrix, b_p an $m \times 1$ vector and x_p an $n \times 1$ vector. (These inequalities typically include nonnegativity restrictions for the variables.) The prevalence of such applications leads us to focus on the linear inequality representation to illustrate our development.

The form of f_p and the coefficients of A_p, b_p may be determined by other parameters. For instance, a person's preference, habit, goals all depend on his or her wealth. They also depend on economic factors such as inflation, interest rates and stock prices. In a financial planning model, for example, a coefficient a_{ij} of A may be defined by:

$$a_{ij} = \sum_{t=1}^{T} \frac{1.0}{(1+\rho)^t} \tag{10.3}$$

indicating the present value of a dollar of income in each period $t = 1, 2, ..., T$. The parameter here is the interest rate ρ.

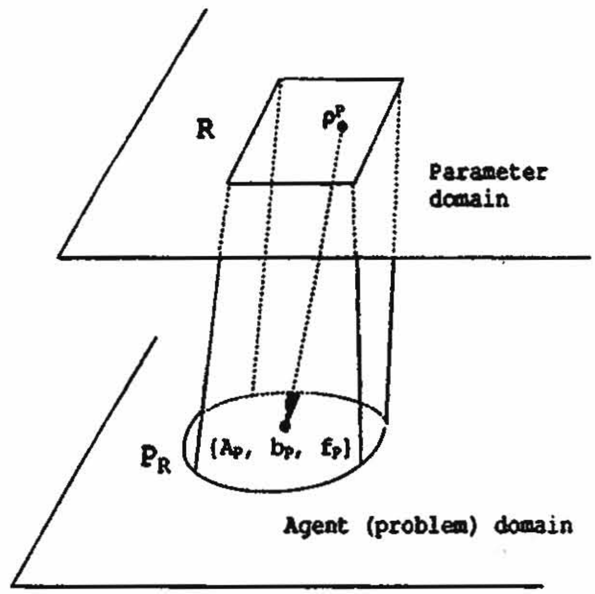

Figure 2 Coefficients are determined by parameters

Formally, we define parameters as data that change independently while coefficients of the optimization problems depend on them. In addition, coefficients that do not depend on any other parameters will themselves be considered parameters. Suppose there are L parameters, $\rho_1, \rho_2, ..., \rho_L$. We stipulate that their special values $\rho_1^p, \rho_2^p, ..., \rho_L^p$ determine the coefficients of A_p, b_p and the form of f_p as follows:

$$a_{ij}^p = h_1^p(\rho_1^p, \rho_2^p, ..., \rho_L^p); \tag{10.4}$$
$$b_i^p = h_2^p(\rho_1^p, \rho_2^p, ..., \rho_L^p); \tag{10.5}$$
$$f_p = h_3^p(\rho_1^p, \rho_2^p, ..., \rho_L^p). \tag{10.6}$$

A triple $\{A, b, f\}$ can be considered as determined by a region in the parameter domain. That is, each instance $\{A_p, b_p, f_p\}$ of the triple is determined by a particular $\rho = \{\rho_1^p, \rho_2^p, ..., \rho_L^p\}$ in the parameter region, as shown in Figure 2. The subset P_R of P identifies an index set for problems whose parameters lie in a region R. That is, we may specify $P_R = \{A(\rho), b(\rho), f(\rho) : \rho \in R\}$. For our present purposes, we will characterize a *region* R of parameters by:

$$\begin{array}{rcl} l_1 \leq & \rho_1 & \leq u_1 \\ l_2 \leq & \rho_2 & \leq u_2 \\ \ldots & . & \ldots \\ l_L \leq & \rho_L & \leq u_L \end{array} \tag{10.7}$$

where $l_i, u_i : i = 1, ..., L$ are lower and upper bounds for the parameters. Parameters and relationships for linear (and nonlinear) optimization problems such as those specified here can be conveniently represented in modeling languages such as AMPL [3] and GAMS [2].

An important task in our analysis is to characterize the population and its subsets in terms of certain features of their associated optimization problems. For example, we may be interested in the maximum, minimum or average of the objective values for all problems in a set $S \subseteq P$, e.g., in:

$$\max\{f_p^* : p \in S\}, \tag{10.8}$$

We define a *grand objective function* for $S \subseteq P$ as:

$$G_{grand}(S) = G(g(p) : p \in S) \tag{10.9}$$

where the function $g(p)$ evaluates characteristics of an individual agent p while $G_{grand}(S)$ evaluates the collective characteristics of all agents in set S. The function g may not necessarily correspond to the original objective function f. For example, it can identify the amount by which the time to solve problem p is reduced if a particular set of elite solutions is selected to provide starting points for an associated set of strategies. The grand objective function can then be a measure of the total time saved (or average time saved, etc.) by such a design applied to a specified collection of problems. When S is determined by a region R of the parameters, namely, $S = P_R$, we write $G_{grand}(S)$ as G_{grand}^R.

2.3 Generalized Sensitivity Analysis

To provide an approach to exploit the foregoing grand objective function effectively, we propose a generalization of the standard notion of sensitivity analysis to allow it to apply to a *set* of optimization problems rather than to a single optimization problem in isolation. Specifically, we seek to analyze how a grand objective function $G_{grand}(S)$ changes as a function of relaxing or tightening a given constraint for all agents in S, or as a function of expanding or contracting a parameter region along one of its dimensions.

For this purpose, given a parameter region R, we introduce a grand dual variable, $D^i_{grand}(P_R)$, which measures the marginal change in a grand objective function G^R_{grand} with respect to a change in a parameter ρ_i:

$$D^i_{grand}(P_R) = \partial G^R_{grand}/\partial \rho_i \tag{10.10}$$

We can use a deterministic simulation procedure to determine the value of this grand dual variable. This yields an approximation, although one may be able to develop an analytic procedure to determine the true value.

We first obtain the current value for a grand objective function in R and then change the range for ρ_i from

$$l_i \leq \quad \rho_i \quad \leq u_i \tag{10.11}$$

to

$$l_i \leq \quad \rho_i \quad \leq u_i + \Delta_i, \tag{10.12}$$

or

$$l_i - \Delta_i \leq \quad \rho_i \quad \leq u_i, \tag{10.13}$$

with the ranges for other parameters of R fixed.

If the change in the grand objective function is Δ_G, an approximation of the grand dual variable in (10.10) is

$$D^i_{grand}(P_R) \approx \Delta_G/\Delta_i \tag{10.14}$$

It is easy to see that our approach reduces to standard sensitivity analysis in the case where we restrict the domain considered. When a region R shrinks to include only one point, the set it generates will include only one agent, or one problem, say p. At this time,

$$G^R_{grand} = g(p) \tag{10.15}$$

which is a conventional objective function for problem p. If furthermore, we are considering a change in the right hand side of problem p, then

$$D^i_{grand} = \partial g(p)/\partial b_i \tag{10.16}$$

which is a conventional dual variable in LP and NLP.

The concept of a grand dual variable is highly useful to characterize the effects of changes in a population. For example, in a setting where a population consists of consumers and the grand objective function measures their purchasing power, alternative grand dual variables can be used to analyze changes in purchasing power that result when consumers grow a little older, or their incomes become higher or their their savings habits change.

Comparing values for a given grand dual variable in two different populations can reveal population differences. For instance, differing degrees of changed purchasing power that result from specific changes in income or savings habits can be pinpointed in this way. Grand sensitivity analysis, by reference to these notions, yields valuable information about factors that influence populations. Issues of investment, product differentiation and product development are clearly affected by such information.

2.4 IPA Cohorts

We say two agents are "associated" if their corresponding optimization problems are related. There are of course many ways to define the condition of being related. For our present purposes we will rely on a definition that implies the problem parameters lie within a certain bounded region, which we embody in the notion of an IPA cohort. In particular, we define an IPA cohort N^a for an agent a to be a subset of P that contains agents associated with a (including a itself) by the following construction. Denote the parameter set that determines a by $\rho^a = (\rho^a_1, \rho^a_2, ..., \rho^a_L)$. Then the IPA cohort N^a is given by a parameter region R^a that arises by specifying lower and upper bounds for the parameters in a manner we subsequently describe. We will restrict attention to IPA cohorts thus generated from such parameter regions and denote them as N^a_R.

The manner in which we determine IPA cohorts will make use of the grand objective function. To motivate this determination, suppose we stipulate that the grand objective $G_{grand}(S)$ is designed to take on larger values when the parameters that define $S \subset P$ lie within greater proximity to each other. Then we may conceive the problem of searching for a "good" IPA cohort for an agent

a to be formulated as seeking a set S that yields an optimal solution to the following problem

$$\max_{S} \quad G_{grand}(S)$$

$$\text{s.t.} \qquad a \in S \tag{10.17}$$

Clearly, an optimal solution results when the set S includes only the agent a, which is not interesting to us. It is therefore necessary for the grand objective function to encompass other factors, which we consider in following sections. When the grand objective function is appropriately defined, the solution to the problem will be N^a.

Since a set of agents and the coefficients for their optimization problems are determined by a region of parameters, which is in turn calculated as in (10.7), the true variables of interest are actually lower and upper bounds of the parameters. Thus, the preceding formulation can be rewritten as:

$$\max \quad G_{grand}(l, u)$$

$$\text{s.t.} \qquad a \in P_R \tag{10.18}$$

$$P_R = P_R(l, u) \tag{10.19}$$

The last equation requires that the region R be determined by $l = (l_1, l_2, ..., l_L)$ and $u = (u_1, u_2, ..., u_L)$.

2.5 Representative Agents

We link the goal of generating a proper form for the IPA cohorts to the goal of generating a collection of *representative agents*, where we require the union of the IPA cohorts of such special agents to include all agents in P. Clearly, by adjusting the bounds that define IPA cohorts, we can assure that any chosen collection of agents will represent the other agents in this fashion, but we base the determination of representative agents on additional criteria.

We are guided by two interrelated concerns. First, we desire the representative agents to be selected so that the agents in an IPA cohort are associated more strongly than simply by sharing a certain proximity to the agent defining the IPA cohort. For this reason, we will in fact allow a slight distortion by permitting "artificial agents" to be included among those we treat as representative, where these artificial agents are derived by perturbing or interpolating between parameter values of other agents. Such artificial agents are generated by starting from original agents and creating adjustments based on feedback about problem solving efficacy from relying on the current representative agents. Alternatively, we can simply shift the role of representative agent from one original agent to another by such an iterative analysis. In either case, the outcome can lead to more appropriate definitions of IPA cohorts for the goals we seek.

Second, once a set of representative agents is (provisionally) specified, we desire to characterize their IPA cohorts so that the agents within a given IPA cohort will be susceptible to exploitation by information that is specific to this cohort. More precisely, we seek to use this cohort-specific information to solve the optimization problems for the agents in the IPA cohort with significantly increased efficiency. Our design incorporates the principles of the learning approach called target analysis [6, 8, 5, 7]. In the present development, as intimated earlier, we choose to rely on information obtained from elite solutions to problems in the IPA cohort. Still more restrictively, we use these solutions to give advanced starting points for members of the IPA cohort. An interesting discussion of the philosophy and potential applications of focusing on starting solutions, viewed from the perspective of case based reasoning, is given in Kraay and Harker [9]. We emphasize that other decision rules and forms of exploitation are possible within the framework of target analysis, particularly in coordination with tabu search as employed in this study (see, e.g., [8]). We have also found that similar ideas can greatly improve the performance of other methods such as in [1, 17]. Nevertheless, we have found this form of exploitation to be highly effective.

2.6 Computational and Practical Benefits

It may appear from our foregoing abbreviated description that we are investing a great deal of effort to solve and re-solve problems, using different initial conditions. This naturally raises the question of how benefit can be gained from such an undertaking, since we are going well beyond the effort required to obtain solutions to the problems in P. The answer is that P is not a fixed set, but rather is one that grows over time, as new agents or agents with new problems enter the scene. (Alternately, in a static "large population" environment, we may consider P to begin as a subset of a substantially larger P that we desire to deal with.) Thus, we are applying the strategy of intensely focusing on problems already solved in order to improve our ability to solve problems yet to be encountered. By expending a large effort to process a relatively small number of problems, we may save considerably greater amounts of effort in solving a much larger number of problems.

2.7 Choice of Agents and IPA Cohort Design

Our construction of IPA cohorts and critical agents is therefore an anticipatory process, where we seek to organize existing knowledge more effectively to meet future problem solving needs. The process is necessarily dynamic and self-modifying. The organization of our knowledge changes as the knowledge changes.

Often initial conditions are simple, and call for simple measures. This is evidently true in the case of choosing representative agents. To begin, we may simply rely on "conventional wisdom" about a problem, according to the context, to select agents considered typical for different ranges of the underlying parameters.

After making such a selection, we seek intervals about the parameters for these initial representatives so that the resulting IPA cohorts will include all agents. However, this goal must be balanced against the competing goal of creating cohorts sufficiently compact that we can identify a relatively small number of useful elite solutions for agents in one cohort – i.e., solutions that yield good advanced starting points for members of the cohort. Upon analyzing the outcomes of selecting various trial elite solutions (by designs we subsequently describe), we may either add, delete, shift or modify the current representative agents to create a better collection.

We may therefore express the overall plan for our approach in the following outline form:

- Choose a list of representative agents;
- Define a proper cohort for each representative agent and characterize it according to a certain purpose. For example, to improve computation efficiency, we need to identify the associated elite solutions. For marketing strategy design and new product development, careful characterization of the cohorts is needed and sensitivity analysis is required.

- If the current IPA cohorts fail to cover all the important problems, select other representative agents and find other cohorts for the new agents.

3 COMPUTATIONAL EFFICIENCY

We focus on applying integrative population analysis to improve computational efficiency. As emphasized, we seek a proper IPA cohort and its corresponding elite solutions for a representative agent in a manner to permit problems in the cohort to be solved easily. On the other hand, cohort size is also a measure of efficiency: if the IPA cohorts are too small, we have to keep many of them to cover a population and must save a large number of elite solutions. In the extreme case, each cohort would contain only a single problem, requiring us to solve all problems in advance with no advantage derived from prior elite starting solutions. When we speak of elite solutions, we refer not simply to solutions that are actually generated for the problems considered, but also to specially constructed trial solutions that are obtained as "centers" of clusters of other solutions. The final determination of a cohort is a compromise between the following two considerations:

1) The average reduction in computation time that results by starting from appropriately selected elite solutions (obtained from representative problems in the cohort) should be as large as possible.

2) Each cohort should cover a large terrain so that only a few cohorts are needed.

An important element in these considerations is that we will simultaneously be seeking to establish a small and effective set of elite solutions for each cohort generated. With this element implicit, we formulate the problem as follows:

$$\max \quad \theta \, TRR_{grand}(N) + (1-\theta)\|N\| \tag{10.20}$$

where:

TRR	:	time reduction rate
$\|N\|$	:	size of cohort N
θ	:	weight to balance the two criteria

The time reduction rate TRR is achieved through a five step simulation process discussed in section 3.1. Briefly, we generate random samples of problems in a cohort, solve them with or without elite solutions and get the average time reduction rate. For each sample p:

$$TRR(p) = 1 - \frac{T_e}{T_{ne}} \tag{10.21}$$

where

$$\begin{aligned} T_e &= \text{computing time with elite solutions for problem } p \\ T_{ne} &= \text{computing time without elite solutions for problem } p \end{aligned}$$

The average time reduction rate in a IPA cohort is therefore:

$$TRR_{grand}(N) = \frac{1}{n_N} \sum_{p \in N} TRR(p), \tag{10.22}$$

where n_N is the number of samples in cohort N.

We can also consider time reduction rate in the constraints, in which case we have the following alternative to (10.20):

$$\max \quad ||N|| \tag{10.23}$$

subject to

$$r_1 \leq \quad TRR_{grand} \quad \leq r_2 \tag{10.24}$$

where r_1 and r_2 are the lowest and the highest time reduction rates which we consider to be satisfactory.

The variable in this formulation is the size of a IPA cohort, which is a function of the ranges of the parameters in (10.7). Therefore, we have:

$$||N|| = N(l_1, u_1, l_2, u_2, ..., l_L, u_L) \tag{10.25}$$

or more specifically,