DEVELOPING SMART AND INTELLIGENT TARGET MARKETING SYSTEMS

Stan Openshaw

WORKING PAPER 94/3

SCHOOL OF GEOGRAPHY • UNIVERSITY OF LEEDS

Developing smart and intelligent target marketing systems

Stan Openshaw, School of Geography, Leeds University, Leeds LS2 9JT

email: stan@geog.leeds.ac.uk
tel: 0532 33 33 20
fax: 0532 33 33 08

1 An emerging new era of computational marketing

This essay presents the outline of what is termed computational marketing. This is regarded as a future essential technological focus based on developing what are cride but nonetheless artificially intelligent marketing systems. It also provides a critical perspective of current marketing uses of computing and modelling methods and makes several suggestions as how both the individual tools and the containing frameworks can be developed into better, smarter, and more intelligent marketing systems better suited for the 21st century.

The concept of database marketing has existed for over two decades. Its origins reflected the increasing importance of computer databases as sources of marketing information. Initially, it was primarily geodemographic codes attached to address lists, then customer lists became available for targeting purposes, and latterly various lifestyle databases. However, there is an increasingly strong argument that these and related response modelling developments are only an intermediate stage in the development of what might be regarded as smart and intelligent marketing systems. The problem with database marketing was that whilst it emphasises databases and computers it failed to generate the tools for optimising the potential benefits offered by either. This was partly a matter of waiting for the computer hardware technology to catch up with the demands created by marketing needs and to provide sufficient compute power to support greater levels of analysis and modelling sophistication. It is argued that this state has now been reached and that the stage is now set for the emergence of new styles of more machine intelligent and smarter computational marketing technologies. Indeed the real challenge today is to develop the new marketing tools that can exploit the opportunities created by these more powerful computing environments. This paper outlines some of the ideas being developed in the School of Geography at Leeds University.

The main justifications for seeking the development of smart and intelligent systems are as follows: (1) the realisation that at present many areas of marketing computing are in terms of their computing positively dumb; (2) current and expected future developments in affordable high performance computer hardware provide an alternative compute intensive strategy for dealing with many marketing problems and now have the capacity to support the development of what are termed smart computationally based marketing systems; and (3) there are major commercial benefits to be gained by developing more intelligent computer technologies relevant to marketing needs. It is noted that at present few marketing applications actually exploit what can be now be done with current computer hardware. This is not really a matter of repeating on a pc or workstation platforms previous mainframe applications. This may reduce costs, but the really exciting opportunities and challenges involve developing completely new applications that previously would either have been unfeasible or impossibly expensive to consider and which start to really exploit the computer as a marketing device by creating new styles of computational marketing systems. It is not just that the economics of computing have changed in a major way but also that the entire computing world is in the midst of a profound revolution with far reaching implications for many activities including marketing.

It is estimated that today it is possible to purchase for about £20,000 a workstation with the same performance of a supercomputer that would have cost 1000 times more only a decade earlier. However, this speed-up of computer hardware and the dramatic decline in costs is by no means over. Workstation speeds will probably increase by up to another factor of 100 during the 1990's whilst prices drop in real terms. At the top end, it is estimated that whilst the speed of the UK's fastest supercomputers have increased by a factor of about 20 during the last decade, it is now confidently expected that there will be a further increase of between 1,000 and 10,000 times in both compute speeds and memory capacity by the late 1990's. Many computer scientists now speak of a new era of computation, one based on massively parallel machines. (Hillis, 1993) and it is extremely unlikely that the marketing world will escape these developments.

Perhaps the most notable feature of this dawning age of virtually unlimited and cheap mass computer power is that many hard or previously impossible to solve problems will start to be overcome via massively compute intensive solutions. This is very relevant to marketing because it constitutes one very obvious highly commercial area in which these developments can be readily exploited to meet existing needs. So, it is suggested that the 1990's will witness the development of a new style of what is termed here computational marketing. Moreover, it might be argued that the principal problem at present is simply that of increasing awareness for what in the early 21st century will be a major outlet for applied non-scientific teracomputing. Whilst there is certainly a need for training and education, there is an even greater opportunity to start to develop entirely new approaches to marketing.

It may have also gone unnoticed that practical and applicable Artificial Intelligence (AI) tools now exist and can be exploited in marketing contexts. Put simply, AI offers the means of adding intelligence to our marketing systems via a computational route. AI is very relevant because it is clear that new methods, new approaches, and new systems are becoming available and it might be expected that some at least will be spectacularly successful. Ten years ago, the computer methods available to marketeers were not qualitatively much different from those in use today. It is true that databases have improved, it is now easy to draw multi-coloured maps and there has been a vast diffusion of computer power. However, by contrast with the last decade, it is now very likely that in ten years time nearly all the world's marketing methods will be based on very different technologies, on smart systems of one kind or another and that the real challenge is not whether it will happen but how best to phase it in. This reflects the view that the 1990s are a time of radical change in many areas of information technology. The long drawn out 'S' shaped innovation diffusion and learning curves of the 1970s and 1980s are starting to become exponential rather than linear. Rates of innovation and change once measured in units of a decade are now being measured in units of 2 or 3 years, and perhaps soon in terms of one year units or less. Such periods of rapid innovation and change are both threatening to the established order and exciting because of the new opportunities they create. The mid-1990s is identified as one such period.

So what is meant by a smart and intelligent marketing system? There are various definitions. Sections 2 to 4 argue that they are the opposite of what is the current practice. Smart intelligent systems have some at least of the following properties:

(1) they learn about their mistakes

(2) they have mechanisms that attempt to combine the best that machines can do with human intuition and knowledge;

(3) they are not expert systems in the conventional sense because there is nothing worse than a supposedly intelligent system that operates in a brain damaged fashion due to fundamentally unresolved problems of knowledge

elucidation, instead a smart system is one informed by human knowledge but not restricted to it;

(4) they are safe, they will self-diagnose failure and provide some assurance of optimality; and

(5) they are adaptive, dynamic, and sensitive to the complex environments in which they have to operate; in some ways they are "alive".

Whilst it is not yet possible to buy "off the shelf" smart and intelligent marketing systems, many of the pieces needed to assemble such systems now exist. However, being smart is not simply a matter of using AI or building better computational tools to replace current dumb methods, but it also involves developing a total systems framework that allows intelligence to exist by providing feedback loops. A smart intelligent marketing system has intelligence at both the local and the global levels.

If this preliminary discussion seems either vague or prosaic then maybe the arguments will become clearer via a brief review of some of the various possible application areas in target marketing. Section 2 looks at geodemographics, Section 3 considers lifestyle databases, Section 4 looks at response modelling and Section 5 focuses on what might be termed global smart and intelligent marketing systems.

2 Developing smart geodemographic systems

2.1 Simple sloppy sixties systems

The geodemographics industry is well established but shows little signs of developing intelligent systems. In fact as it is currently practised, geodemographics is a dumb targeting technology. It is largely based on an unfortunate mix of: (i) old fashioned methods that now need updating; (ii) a neglect of almost everything that is already known about the characteristics of residential neighbourhoods; and (iii) is applied in a deterministic fashion without any opportunity for much input of system based intelligence. Table 1 outlines some of the principal sources of dumbness.

This dumbness is sad because in many ways geodemographics is a future essential targeting technology; principally, because it is and is likely to remain Data Protection Act proof and because it does not make as many unwarranted assumptions as do more individual person based forms of targeting. The problem is that the geodemographic concept has not been developed into the real neighbourhood marketing tool that it is often claimed to be. In fact, viewed from a technological perspective, it is still largely based on what might be derided as the triple 'S' concept; Simple, Sloppy, Sixties systems. Geodemographics needs to be re-invented as a 1990s technology. Three areas need attention: the classification process, the incorporation of knowledge, and the development of sensible and intelligent targeting tools.

The 1991 census seems mainly to have stimulated an understandable desire to repeat the 1981 census based systems with 1991 data. However, few of the lessons have seemingly been learnt; nor have the dramatic changes in the economics and power of computing been much reflected in the development of better and more intelligent geodemographic systems. However, a few trends are evident: a greater emphasis on customisation, on supplementing census data with other sources of marketing relevant information, and linkages with Geographic Information Systems (GIS) (Openshaw, 1993). GIS provides a nice gloss on a technology that desperately needs to be redeveloped rather than simply be re-presented in a not particularly helpful map based context. GIS is relevant technology but it needs to be used to improve geodemographic targeting rather than merely make it more colourful. The current enthusiasm for GIS in Business tends to overlook this severe structural deficiency.

The great strength of geodemographics is its nature as a small area targeting tool. It is a shotgun rather than a sniper's rifle. Yet there are a number of ways it can be improved; these include: (1) the use of modern rather than obsolete classification methods that handle rather than ignore the problems of the data being classified; (2) tuning the classification to meet particular market sub-sector or specific client needs; (3) use of the most marketing relevant information, which may well not be census based; (4) the use of up-to-date rather than old information even if this means a move away from census based geodemographic systems to data sources that can be updated on an annual basis; and (5) by developing appropriate and smart targeting technologies that sit on top of the geodemographics to replace the simple client profile based targeting that is traditionally used in order to exploit 'real' neighbourhood effects.

It is noted that none of the standard classification procedures found in widely available statistical packages or indeed those traditionally used in a geodemographic context are really suitable for creating geodemographic systems for the 1990's; for example, this would include the CCP software used to create Super Profiles 1981 (Charlton et al, 1985) and still among the best of the standard methods today. Ten years ago, there was little or no choice; but this is not true anymore; see Openshaw and Wymer (1994). Yet, the more refined the classification attempts to become the greater the problems. For example, mixing census and noncensus data might seem to be a good idea but it causes tremendous methodological difficulties. The census data are reported only for enumeration districts. When these data are merged with noncensus data for unit postcodes, then there are massive problems of differential data accuracy which conventional methods ignore. The results may appear more relevant (being postcode based) but the quality of the classification might easily have been reduced. Small number problems and varying levels of data accuracy, precision, and reliability need to be handled rather than overlooked. Unless considerable care is taken, then the apparently more refined geodemographic systems might actually be poorer than the much cruder attempts. It is really a matter of balancing the errors and uncertainties in the various parts of the geodemographic targeting process. Too much accuracy in one part may not be translated into better results.

2.2 Improved geodemographic classifiers

One key component is the residential neighbourhood classification and it is here where the greatest attention needs to be initially focused. Fortunately there have been major progress recently in the development of new types of classifier that seem particularly relevant to handling many of the problems associated with small area census and marketing data (Openshaw, 1994). In particular, neural net based methods can be used to classify such noisy data, provided you can afford the computer time.

The question of classification tuning also needs to be addressed. General purpose geodemographic systems and even sub-sector systems are really only a targeting tool for the desperate. If you have no other data and are interested only in quick and dirty old-fashioned targeting tools then maybe you have no choice. However, if you are not in this category, then you should today be seeking customised and tuned geodemographics. Why should you be using the same segmentation system as your rivals? What benefits are there in repeatedly hitting the same old tired target groups? Does it really make sense to have a single system for the whole of the Financial Services sector? Equally, if you want to customise, then what additional benefits are there if you based your customisation on old fashioned and obsolete classification methods. There is just no point in trying to tune a model T ford for a modern Grand Prix. Customisation will probably only work well when the classification methods are sensitive to the data and clever enough to handle rather than ignore the problems that a higher degree of focusing creates.

Indeed the more you wish to customise; and perhaps include your own data in the exercise; the better the classification methods have to be otherwise you may not gain much or any extra targeting precision because of the problems of handling data of mixed and variable levels of resolution.

This discussion illustrates another key point. The falling cost of computing and the speed-up in computing hardware over the last decade has trivialised the task of creating geodemographic systems. These systems can now be created on a workstation or a pc in a few hours or a weekend. Back in 1984-5, the 1981 Super Profiles system required something like 24 hours of CPU time on what was then a very large, fast, mainframe computer. Now it can be quickly recreated in a few hours on a workstation, but who would really want to do that? It is true that the classification methodology used for Super Profiles is still being widely applied, indeed it is still better or equivalent to many good methods, but it is no longer state of the art. The speed-up in computer hardware should be used to develop even better classifications. Sadly, there is little or no evidence that this is happening; this is a great mistake that can only reduce the power of geodemographic systems. The challenge is not to reproduce many new Super Profile or ACORN or MOSAIC like systems for many different clients; but to apply much more sophisticated and data relevant methods so that the overall quality of the classifications are substantially and significantly improved. Currently, one suspects that the principal benefits of the falling cost of computing is not being passed on to the users.

As a result it is likely that most 1991 based geodemographic systems will work no better than those of a decade earlier and a major opportunity for progress will have been lost.

2.3 Incorporating geographic knowledge

Another key aspect of the classification process causes concern. Virtually all geodemographic systems so far produced assume that nothing is known about the data being classified. The task of uncovering structure is left entirely to the computer based classifier. Human knowledge is only of marginal importance; it is used to label the results and, perhaps, to determine a useful number of clusters. This is a gross neglect of a half century of research on the socio-spatial structure found in cities. The knowledge exists in the literature but it is currently not used to guide the classification process. As a result, the geodemographic classifications are conceptually weak and, perhaps, worse there is no logical consistency checking; for example, do urban residential area types only appear in urban areas and not in rural areas, are concepts of social gradients within cities preserved or destroyed, do the results make sense locally as well as globally? Without the input of geographic knowledge there is no guarantee that the results are going to be sensible. This is a severe criticism. Fortunately, it can now be handled and better residential area classifications obtained.

Figure 1 illustrates one way of operationalising geographic knowledge. The areas which are identified as being potential errors are re-submitted to the classifier to see whether it can be re-assigned to another cluster that is better (in a global sense) or to a nearby cluster that improves the knowledge "fit" (in a local sense) even at the expense of a minuscule loss of global performance.

2.4 Smart geodemographic targeting

A final aspect that is relevant here, concerns the further development of geodemographic targeting. Current systems are not intelligent. It is now trivially easy to put postcode level geodemographics on pc platforms. A complete system for a 486 pc requires about 10 MB of disk space and is extremely fast. However, this is no better than what was previously provided by mainframes; apart from possibly it now sits on your desk and has a mapping capability. The results will be no better than a decade previously as there is no intelligence other than that which the user provides. It is argued that this is no longer good enough.

Why not use the available computer technology to add a new dimension to the targeting. Openshaw (1989, 1989b) describes how to build a fuzzy targeting procedure. Until recently this was infeasible as a practicable and portable tool, mainly because of computing limitations. These constraints are no longer relevant. The method involves looking at uncertainty in both the "geo" and the "demo" parts of the geodemographic process. It works as follows.

It should be well known that if a target market consists of A1, A2, A3, and A4 types of areas and you mail the people living there, then you would get a range of responders, not all (or even most) would have typical A1-A4 characteristics. This is because the residential area types in all geodemographic systems are composed of a mix of different features. They are not homogeneous even if the trendy labels or "thumb nail" sketches suggest that they are. For example, even highly distinctive wealthy or ethnic or old-aged area types will often have a majority (sometimes a large majority) of people with other characteristics not well represented by the area labelling; see Openshaw (1984). This reflects the nature of census data, the problems of spatial classification, the appropriateness of the classifier, and errors in the generalisation of all of Britain's residential mosaic into a relatively small number of types. It follows that postcodes allocated to A1-A4 types will vary in their possession of these characteristics. More seriously from a target marketing perspective, it is likely that many postcodes assigned to non A1-A4 may well have some or many of the desired characteristics. It is really dumb then to just target A1-A4 areas as there are errors of inclusion and exclusion.

The "geo" part of geodemographics can also be improved. Presently, there is still some uncertainty as to which residential neighbourhood type a postcodes belongs to. Only when the Ordnance Survey's Address Point product is completed in 1995 will this residual uncertainty largely disappear. However, this ignores an even more important point. Current supposedly neighbourhood classifications have no explicit neighbourhood effects within them! There is, however, no reason why explicit neighbourhood effects cannot be introduced. Neighbourhood effects are the notion that people who live "close" or "near" each other will have aspects of their behaviour in common regardless of their geodemographic characteristics. This is a well known phenomenon and is implicit in geodemographic targeting which makes use of neighbourhood effects that only that part which operates within the census enumeration district, but not across its boundaries. Incidentally, this effect is much weaker at the unit postcode scale.

A geodemographic targeting system should be devised then to target households not in A1-A4 residential areas but very near or geographically contiguous to them. It is here where any real spatial neighbourhood effects may be found. Neighbours in the same street, neighbours close in their geographic proximity, and neighbours within line of sight of each other, are all important if purchasing behaviour is really the spatially contagious process that many believe it is.

Fuzzy geodemographics attempted to combine both sources of uncertainty to develop a more intelligent geodemographic targeting process. At the very least it involves adding extra dimensions to the tradition set of neighbourhood types. Figure 2 illustrates what is involved. It increases the size of the associated geodemographic database to about 100MB. This is not a major problem anymore, although the real clever bit is how to utilise the extra dimensions added to the data. One method is to use the results from a mailing to determine appropriate levels of fuzzyness in both the geographic and demographic dimensions for subsequent mailings. Another is to use profiled databases to develop a characterisation of target groups that include in varying degrees the geo and demo fuzziness appropriate for them. These are potentially very useful and exciting developments to a targeting technology that has been unnecessarily dumb for far too long.

Finally, you need to hide the complexity of fuzzy geodemographics behind an appropriate user interface. It is very important to retain the ease of use properties of geodemographics and this can be achieved by keeping the user interface simple, by automating the fuzzy calibration process, and then building in performance monitoring and self-evaluation tools. It is also most important, as far as is practicable, to build geodemographic systems that have feedback based on performance built in from the beginning.

3. Developing smart lifestyle and client information systems

3.1 Data overloaded lifestyle marketing systems

One reaction to the dumbness of geodemographics has been to switch attention to customer orientated marketing based on databases consisting of large amounts of personal data with many seemingly marketing relevant variables. The developments of computing hardware over the last few years has made this possible, economic, and increasingly make it easier to hold vast online individual databases, although seemingly the secret of building good customer databases still eludes many large organisations. Table 2 outlines the principal areas of dumbness applicable here.

The main problem now is how to make good, appropriate, and full use of the vast quantities of data that are now available. If the targeting part of geodemographics seems crude, then what goes on here is really no smarter. Typically, the user is supposed to know the characteristics of the potential clients that are to be mailed. It is on this information that the performance and efficiency of the entire approach largely depends. The sniper's rifle only really works if the sniper is a good shot, knows precisely who to aim at, and actually hits the right micro-person. There is now very little scope for error. Unlike geodemographics, this is not fault tolerant technology. Various statistical response modellers can be used to assist in the selection process; although as the next section argues these are themselves often fairly dumb although usually due to no fault of their own making. Lifestyle mapping systems may help provide a platform for developing a better understanding of what is happening but this is clearly not by itself sufficient as a useful targeting tool as distinct from being a useful source of background information.

The principal problem is that most of the information contained in an individual level databases will not used. New methods are urgently needed to "mine" such databases for marketing purposes. Additionally, it is important to try and close the system loop by building an intelligent and adaptive system. All too often, valuable response information is seldom collected, treasured, saved, and then processed as a means of improving selection efficiency. Database marketing is often being practised in an uniquely dumb manner with little scope for smart behaviour other than what can be input by the skills of the marketeer. As databases become larger and more complex, as markets become more confused and customer orientated, as people's behaviour becomes less predictable, as instabilities and uncertainties in the enclosing social and economic environments continue to increase, so the need for inherently smart technology is rapidly growing. The marketing process needs to be viewed as a single integrated system even if as at present many of the component subsystems are separate and unconnected entities that are operationalised via a number of different, often remotely located agencies, who operate without much or any feedback, and seemingly without any global systems view or control.

3.2 Many databases are dynamic living systems

The discussions in section 4 are very relevant here but so too is the view that more effort should be put into recognising that lifestyle databases, and even more so many client and customer databases are living entities. Fresh information is constantly being added. There are many, almost continuously occurring, opportunities to evaluate and test models relevant to database marketing. New data is far too often, if not always, regarded as a database management task (ie update or replace or delete or deduplicate). What a waste! Instead of replacing and destroying or merely adding, there is an opportunity to test and improve the performance of database selection rules or response models. The database at time t can be used to predict time t+1 and then later compared with the observed t+1 data. In this way database development simultaneously becomes a continuous test of targeting methods; if the targeting methods are smart they will improve their performances with experience over time, and in the process become much more valuable.

Typical questions that users currently expect lifestyle databases to answer include: (1) how to find more customers like these customers; (2) what idealised profiles do good customers possess; and (3) how best to target low response but high value prospects. Instead of answering these questions directly, the systems will typically auto-generate a range of frequency histograms from which the user is somehow supposed to deduce something useful. Quite often it seems the marketeer is lured into performing an uniquely dumb mail-out. The selections are made a priori by intelligent users who had better be good, because there is no relevant feed-back. Moreover, the development of optimal and intelligent best performing targeting methods is not usually in the list brokers or database vendor's best short-term interests. Their profits are a function of gross volume of names mailed; indeed, methods that might achieve better response for a greatly reduced mailout may in the short-term be considered very undesirable by all concerned, although if costs are related to targeting performance and then bulk volume, such technology might well be more readily accepted.

3.3 Mining databases

It is noted that AI tools that could be used to generate optimal database selections exist and are fairly easily applied. Two examples help to illustrate this point. The last five years have seen much progress in machine learning tools that can be used to deduce rules for selecting high probability responders from a database, when provided with feedback or training information. Neural nets provide one way of attaining this goal. However, another extremely effective method is a variant known as exemplar learning.

Imagine a database of 250,000 records, profiled in the usual way. There are 1,434 responders. The task is to use the first 50,000 records which contained 151 responses to predict the remainder. An exemplar learning method correctly predicted that 197,000 would be non-responders, it correctly identified 436 responders and missed 690. However, it would only have identified a few thousand clients to mail, of which 436 responded. Of course, to effectively exploit this method it would be necessary to either save historic response information for subsequent training purposes or use a sequential mailout process with feedback. It is also likely that performance would be application dependent.

Another class of database relevant AI technology is the Artificial Life literature; (Langton 1989; Langton et al 1991). The concept is easily described but its implementation is left as an exercise for the reader; see Openshaw (1993). Imagine a flock of artificial "creatures" that are free to roam around the database. Their objective is to identify idealised client profiles (in terms of the available variables) that when used to select real clients will select those with a high propensity to either respond (or refuse). The creatures "breed" using technology found in genetic algorithms (Holland, 1974; Goldberg, 1989). Their behaviour is: bottom-up and emergent rather than fixed or predetermined, they are highly flexible, assumption free, and adaptive autonomous agents. Their sole raison d'etre is to 'find' good performing client profiles. These creatures can cope with the complex geocyberspaces that are databases. They are not blinded by complexity nor limited by our ignorance of what to expect. It is possible to design creatures customised to particular purposes. They are smart because they can adapt, to changes that matter without being instructed to do so; ignoring those that do not. The full benefits of smart techniques is, however, only going to be obtained in a fully intelligent systems framework.

4 Developing smarter response modelling technologies

4.1 Old fashioned static response modelling

Modelling is in principle inherently intelligent technology. An attempt is being made to represent and predict some system of interest usually in an optimal statistical sense. Indeed it is now possible to build highly sophisticated models and there is an extensive literature that might be applicable to marketing. The problem here is that a poor model might well be worse than no model at all. Indeed as the range of available response modelling methods become greater and more sophisticated so the problems of insecure performance may be increasing. Table 3 summarises some of the principal causes of dumbness associated with current response modelling methods.

Improvements here are dependent on two key parameters: (1) the nature of the methods themselves and their ability to handle the data and applications that characterise target marketing; and (2) on the development of a systems view of the wider macro marketing process so that there is feedback and an opportunity for adaptive behaviour to occur. Increasingly, it is this latter factor that is most important. It seems that even a fairly dumb, static, simple response, model can be revitalised by embedding it in a dynamic feed-back situation. It should be fairly evident that the world is a highly dynamic place. Consumer behaviour is not static, nor it is that well understood or well modelled. The decision to consume is largely but not completely a random process. Even the best current models will have percentage error rates of between 90 and 99 percent in marketing applications. Good models may fail or perform poorly for reasons which are not apparent or understood or could have been predicted K weeks (or days) in advance. Its a hard world! So being massively dumb and static is not a good idea if there is any prospect whatsoever of being slightly cleverer.

4.2 Neurocomputing models

There is no doubt that the technology used for response modelling can be improved. Neural nets provide the basis for a highly flexible new technology. They provide the ultimate in black box modelling suitable for a wide range of purposes and offer a means of non-parametric response modelling. Whether they out perform more traditional techniques depends partly on the application and partly on the statistical skills of the user. However, this may well be an increasing irrelevant comparison. If a neural net in the hands of an unskilled user can even get close to the performance achieved after a lengthy careful, handcrafted, highly skilled and expensive modelling exercise; then that may well be more than acceptable. The principle issue is rapidly becoming that of choosing between alternative methods and how to obtain the best and safest possible result. This applies to all forms of response modelling and not just the neural net sort.

The basic problem is the inherent difficulty of the response modelling task. If a statistical method is used, then it is quite likely that some if not most of the key assumptions will not be met by the data. Typical response rates are often too small for reliable parameter estimation. Major issues of good model design and specification have not been adequately resolved, and there are features of the data that render it hard to analyse, particularly small number effects, high but variable noise and error levels, and a high degree of multivariateness. Clearly, it is essential to make good use of those statistical methods and models that do exist, whilst putting more and more emphasis on careful cross validation and testing.

Neurocomputing tools are increasingly being used to model response; for example Furness (1992), Murray (1993). The technology is highly attractive, even infectious in its appeal. There are likely to be many new applications in the next few years as the necessary software is already widely available and often free. It is also an extremely useful technology that handles many of the problems that conventional response models find so difficult. It is, indeed, likely that black box response models of the neural net type will take over. The problem is increasingly how to avoid Openshaw's law (Openshaw, 1992) which states that "the greater the sophistication of a targeting method or response model the greater the potential performance gain but equally the greater the risk of obtaining totally useless results". This principle is based on experience in marketing and reflects the view that whilst conventional, berated as old fashioned methods, are often mediocre or poor performers, they are seldom either very good or disastrously bad. By comparison, more sophisticated methods are more likely to be extreme in their behaviour; either very good or extremely bad. The improvement in performance is often gained at the expense of lost robustness, or at least a reduction in it. It is increasingly important to discover how to make safe and optimal use of response models, particularly as the range of models increases and the technology becomes less understandable. Safety will become an increasingly important issue.

4.3 Marketing Machine 1

One useful suggestion here is to develop a fully automated response modelling based Marketing Machine; the MM/1. This might consist of a set of several different methods that can all be applied to target marketing problems; see Figure 3. The intelligence comes from automating the technology and then selecting the best results having run through a number of different methods; for example, geodemographics, step-wise regression, logistic regression, discriminant analysis, chaid and a few neural nets. The smart bit comes from the evaluation and validation of the models. The user has to decide the trade off between performance and risk and not be too concerned about the modelling technology. The user can select the criteria to be used; some of the principal alternatives are: maximum data fit, minimum risk, maximum profit, or some user specified function. So the "smartness" in MM/1 results from the evaluation of alternatives, reducing the end-users need to choose in advance a global best method whilst providing some assurance that for any particular application, a good or even best model has been found that produces "safe" results.

The MM/1 concept offers the unique theoretical benefit that it should be impossible to easily obtain significantly better results. As new methods are developed they can be added to the list of models included in the evaluation. The strategy is simply to use computer time as a substitute for high levels of human modelling skills and also to remove the uncertainty caused by an increase in the number of alternative models available. The user no longer has to choose a single method. The MM/1 is feasible because of the falling cost of computing and the speed-up of hardware. It is sensible because it minimises the risks of coping with new technologies; if the newer methods work better than old favourites and are safe, then they will be selected and used. Clearly MM/1 will not always be appropriate but there are many applications when it will.

4.4 Marketing Machine 2

The MM/1 concept can be developed further. Instead of forcing a choice decision to be made between different sets, a better strategy in some circumstances maybe to seek an optimal combination of model predictions based on experience. Figure 4 outlines a simple learning response modelling system. A neural net is used to take the results of k different methods and provide an optimal weighting, trained on historical data for a particular type of response application. It is noted that many marketeers have a wealth of historical response data that is valuable if it can be used to improve subsequent mailouts. In figure 4 these data are used to train a neural net to combine two or more different model predictions. The system will learn how to improve itself as well as adapt to sudden or unexpected change. The system also fine tunes to a particular situation and whilst this may reduce generalisation properties it might also be considered a major application specific advantage, once the training data set exceeds a certain size.

4.5 Genetic programming models

One further possibility is to use a less brittle response model. The model breeder machine of Openshaw (1988) can be used to create robust models. This Automated Modelling System (AMS) prototype was subsequently converted into the OMEGA (Barrow,

1993) product. Clearly, the basic principles of symbolic response modellers created by a genetic algorithm to provide a robust best fit to response data can and will be developed further, utilising developments in genetic programming; see for example Koza (1992). Whether this completely negates the dangers of brittleness is probably doubtful, and maybe the AMS/OMEGA approach should be viewed as yet another form of nonparametric regression, serving a function not unlike that provided by a neural net. The development of optimal performance robust models might well in the longer term provide an alternative to the MM/1 and MM/2 concepts. The models are optimal in that they directly seek to optimise the user's expressed needs; for example best fit or least dangerous or maximum profit and low risk. The performance now depends not on the pool of alternative models or their optimal local combination but on the ability of the model breeder to find appropriate models from a much larger universe of all possible models that could in theory be generated from the available pieces. This general goal may only be a few years away from achievement; and, it is certainly worth experimenting with now in those application areas where good response modelling is hard but absolutely essential.

5 Developing smart and intelligent systems

A key feature of smart and intelligent systems is an ability to learn from mistakes. A dumb system having made a mistake once simply continues to make it; or having discovered something useful has no means of discovering this fact or refusing it; or more commonly fails to realise it has made a mistake and therefore never improves. In a marketing context, if you select the wrong target group then you may well obtain poor results. Many marketeers respond either by cunningly not monitoring performance or by blaming someone else; others imagine that it is impossible or hard to any better. The whole task of managing failure is made easier because failure is a relative term and there is no good or obvious way of knowing what could have been achieved. Moreover, everyone is presumed to be in a similar situation. Being intelligent is not just a matter of discovering how to manufacture in a smart or non smart way better selection statements for a database, or of developing better and smarter response models. Whether the source of the intelligence is the user or some software is also fairly irrelevant. There is no need to choose, you use both. However, computational marketing systems are not just smart they are also applied in an intelligent way. There is a need to link the personal details and lifestyle information of clients with the detailed nature of the offer and its packaging in an optimal customer relevant fashion. The solution is to extend the adaptive modelling to include all aspects of the marketing process where there is an element of choice. These areas include: the creative aspects of packaging, the detailed nature of the offer, and the selection of names to be mailed. Once the areas of choice have been parameterised, performance of the entire system can be optimised in an adaptive fashion using AI methods.

The recommended approach is to develop marketing system that contain inbuilt performance feed-back loops. Ideally, these loops should operate quickly enough to allow the modelled selection and decision making processes to be changed in the light of experience; perhaps in near real-time. At its very simplest, a model or selection rule would be regularly re-run to capitalise on the results of a sequential or batch orientated mailing campaign. Better still would be the utilisation of an inherently intelligent model that possessed memory of both past successes and failures and which could quickly respond to changes in the underlying but unknown consumer behaviour patterns. It matters little that current levels of process understanding and causality are poor; because the database of responses contains much of the useful information; if only it could be fully used. Appropriate methods do exist and can be extremely effective when embedded in a sequential mailing process. The intelligence comes from being able to respond to the unexpected and without being told exactly what to do next. The dumbness in current systems comes from their static nature, the all at once selection process, and perhaps sometimes also from simple minded arrogance, that the behaviour of the consumer is sufficiently well understood that nothing else need be done.

5.2 Biological marketing systems

It is noted, for instance, that the printing technology that is now available offers a capability for letter customisation far greater than that currently used. For example, it is not unusual to offer customised messages to clients based on look-up tables. This is applied in an a fixed priori deterministic fashion. The print house is separate from the targeting house and the response goes somewhere else. There is no total systems view. However, there is no reason why the selection of names to be mailed and the content of the customised message cannot be dynamically changed under the control of an intelligent controller (ie response modeller) such that overall performance increases as feedback and learning proceeds via a marketing experiment performed on real people. The controller not only makes selections from within a global set but also varies the nature of the offer albeit within predetermined limits. There is here the faintest glimmerings of an intelligent and smart marketing system. Human inputs and skills are still needed. The secret though is to combine what human do well with what intelligent computer tools do well.

A further development might be to use a genetic algorithm as the control process and thus create the world's first true biological marketing system. This could work as follows. It is assumed that all significant, response critical, choices relevant to a particular marketing campaign can be parameterised and then optimised in a highly efficient global manner. The offer has to be parameterised; for example, the costs of a health insurance might be variable. The selection of names to be mailed is parameterised. The range of different personalised message contents and various other creative alternatives can also be parameterised. A large but small global random sample of mailings are made. The response is analysed and the various design parameters changed under the control of a genetic algorithm. Response performance might be expected to quickly improve over time. The entire campaign is sensitized to any relevant changes in the marketing environment. Safety can be enhanced by the simple expediency of starting the genetic algorithm from good rather than completely random sets of design parameters. Maybe with a little luck and development, this generic style of widely applicable intelligent marketing system could be adapted to sell anything. The human skills and inputs are still very important but the sections that human have never been good at would now be done by computer. The system would self-optimise and be responsive to changing marketing conditions without the user having first to be aware of them. It is useful because most marketeers are being blinded and increasingly becoming blinder to many of the important patterns and relationships that are contained in the masses of data now being stored in their databases.

6 Conclusions

The present is an exciting time to be thinking about the possibilities of developing new computational technologies for marketing. Over the last few years, there have been major increases in the amount of marketing data that are available; there has been a major revolution in the handling of geographic information, in the availability of digital map databases, and in the development of GIS based spatial decision support systems. There have also been major changes in computing hardware, and there is increasing evidence of the beginnings of exciting developments in the application of marketing relevant Artificial Intelligence methods. However, in many ways the real computing revolution in most areas of marketing is still awaited. Yet it is now likely that during the second half of the 1990's major new styles and developments of computer dependent marketing systems will appear. Smart and intelligent entirely computer based marketing systems will begin to provide a basis for an entirely computational AI approach to marketing with the marketing process increasingly viewed as a single, integrated, intelligent, albeit heterogeneous and distributed system. As computer speeds rapidly increase, as costs rapidly diminish, as even the largest databases become easier to handle, there dawns a new era in marketing. This process is now well underway with new products, new sources, new approaches all becoming computationally based. New smart and intelligent marketing systems are in the process of being created by essentially throwing computing power at the problems. We become smarter by becoming computational and we become intelligent by building in feedback and continuous adaptivity. There are immense potential opportunities for those organisations who seek to exploit these developments and, equally, there are immense dangers for those who fall too far behind.

Acknowledgements. Thank you GMAP for permission to talk about some of the computational marketing technologies that are currently under development there. Some of the biological marketing ideas were stimulated by conversations with Neal Bradman.

Table 1 **Dumb geodemographic systems**

based mainly on census data
use on old data
poor old fashioned classifiers
general purpose segmentation objective
improperly customised systems
simplistic targetters
static
not updateable

Table 2 **Dumb lifestyle databases**

selections left to user
too much data to comprehend
quality an unknown quantity
progressive list exhaustion
unrealistic user expectations
no feedback
data protection complications
expensive?

Table 3 **Dumb response modelling**

too statistical
too simple with global functions
neural net hype
small response causes problems
hand crafted, highly skilled and expensive
no feedback
non-adaptive methods

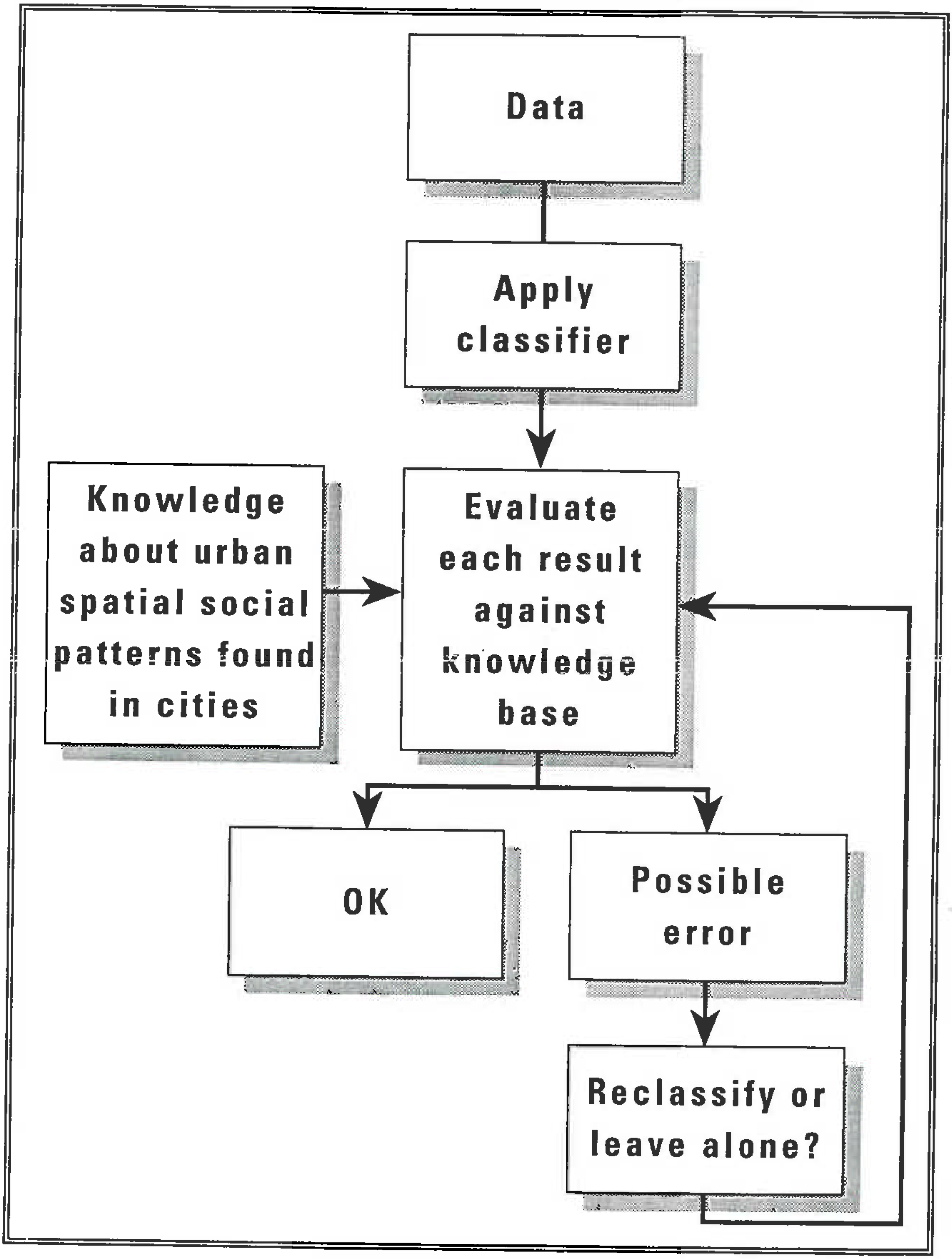

Figure 1: A Knowledge Based Classification of Residential Areas

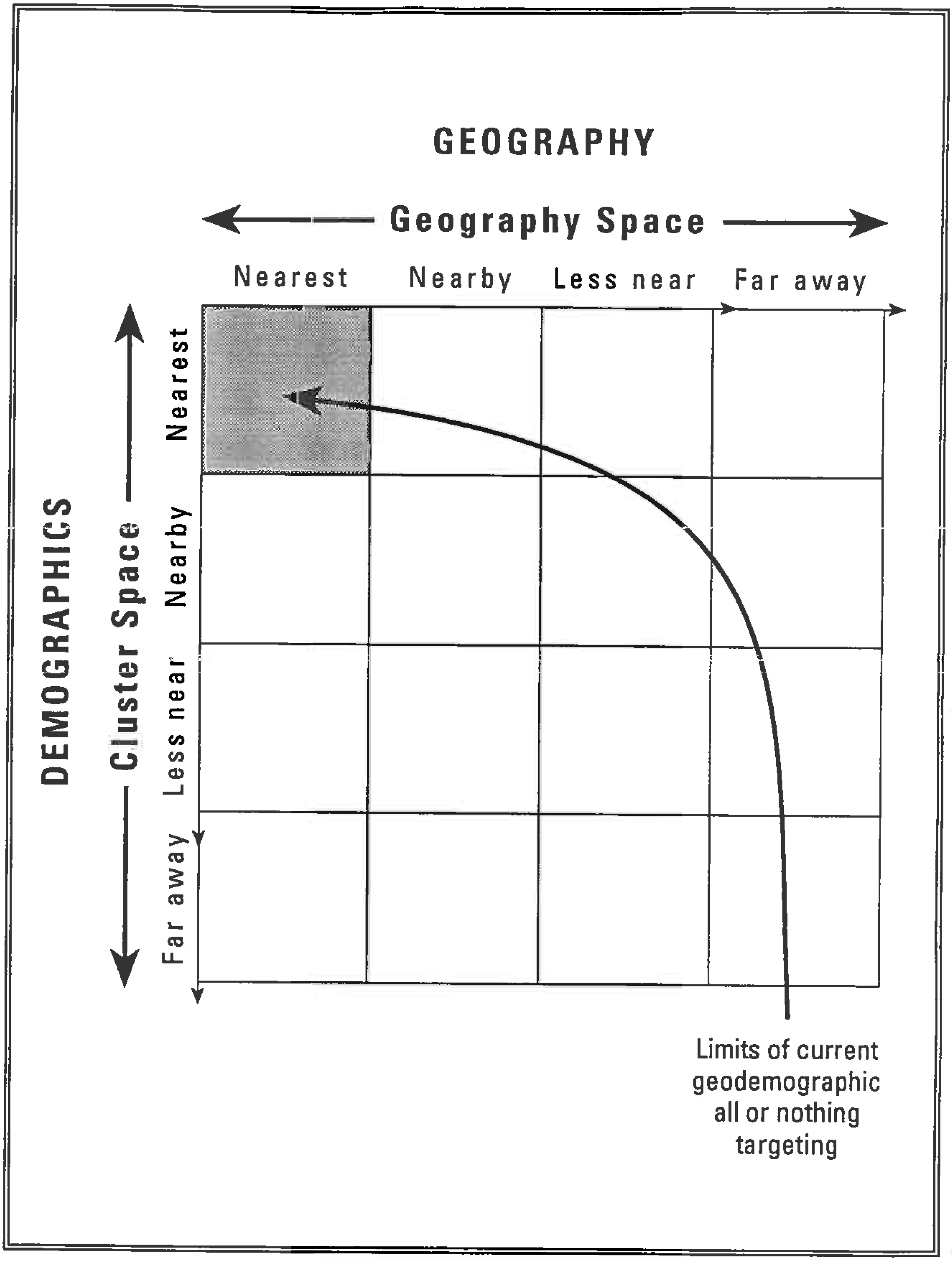

Figure 2: Fuzzy Geodemographics

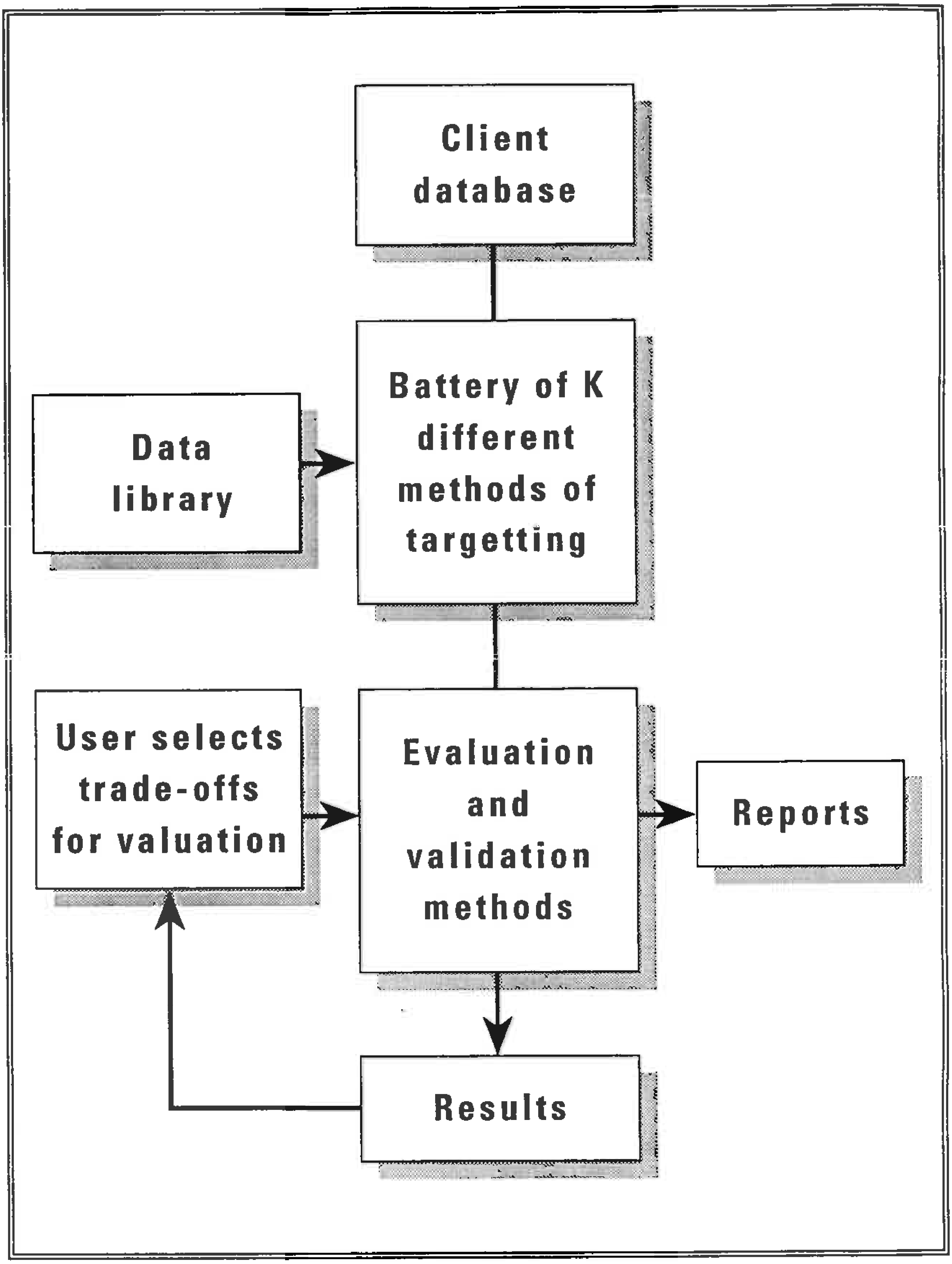

Figure 3: MM/1 Concept

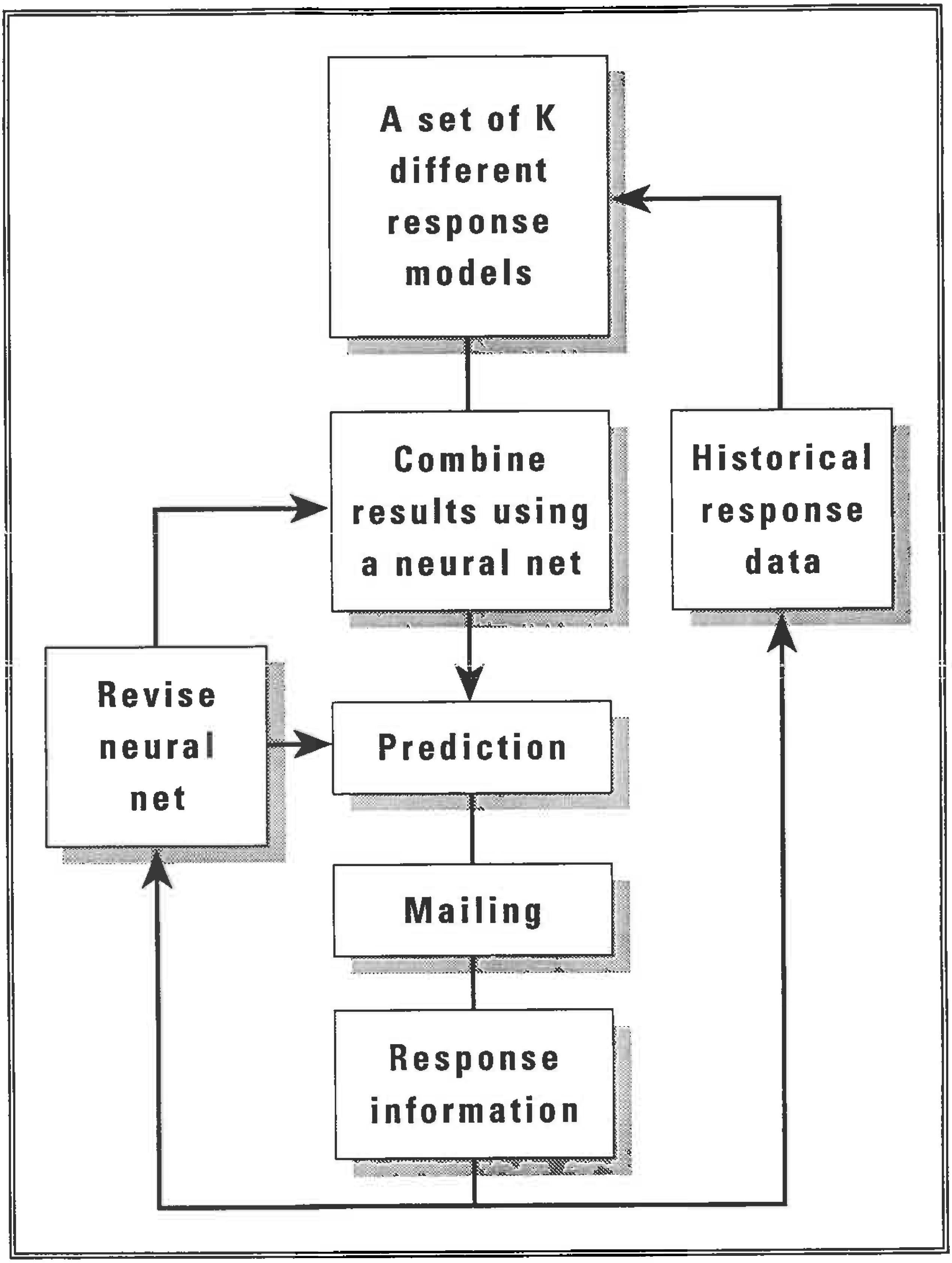

Figure 4: MM/2: A Learning Response Modeller

References

Barrow, D., 1993, 'The use and application of genetic algorithms', J of Targeting, Measurement, and Analysis for Marketing 2, 30-41

Charlton, M., Openshaw, S., Wymer, C., 1985, 'Some new classifications of census enumeration districts in Britain: a poor man's ACORN', J. of Economic and Social Measurement 13, 69-98

Furness, P., 1992, 'Applying neural networks in database marketing: a review', J. of Targeting, Measurement and Analysis for Marketing, 1, 152-169

Goldberg, D.E., 1989, Genetic Algorithms in Search, Optimisation, and Machine Learning, Addison Wesley, Mass

Hillis, W.D., 1992, 'What is massively parallel computing and why is it important?' Daedalus Journal of the American Academy of Arts and Sciences, Winter, 1-16

Holland, J.H., 1974, 'Adaptation in Natural and Artificial Systems', Ann Arbor: The Univ of Michigan Press

Koza, J.R., 1992; Genetic Programming MIT Press, Mass

Langton, C.G., 1989, Artificial Life Addison-Wesley, Mass

Langton, C.G., Taylor, C., Farmer, J.D., Rasmussen, S., 1992, Artificial Life II, Addison Wesley, Mass

Murray, J., 1993, 'Experiences of neural networks in practice' J. of Targeting etc 2, 23-29

Openshaw, S., 1984, 'Ecological fallacies and the analysis of areal census data', Environment and Planning A 16, 17-31

Openshaw, S., 1988, 'Building an automated modelling system to explore the universe of spatial interaction models', Geographical Analysis 20, 31-46

Openshaw, S., 1989b, 'Making geodemographics more sophisticated', J. of the Market Research Society' 31, 111-131

Openshaw, S., 1989, 'Learning to live with errors in spatial databases', in M. Goodchild and S Gopal (eds) The Accuracy of Spatial Databases Taylor and Francis, London 264-276

Openshaw, S., 1992, 'A review of the opportunities and problems in applying neurocomputing methods to marketing applications' J. of Targeting, Measurement and Analysis for Marketing 1, 170-186

Openshaw, S., 1992, 'Some suggestions concerning the development of AI tools for spatial modelling and analysis in GIS', Annals of Regional Science 26, 35-51

Openshaw, S., 1993, 'Two exploratory space-time-attribute pattern analysers relevant to GIS', in S Fotheringham, P Rogerson (eds) Spatial Analysis and GIS Taylor and Frances, London p83-104.

Openshaw, S., Wymer, C., 1994, 'Classification and regionalisation' in S Openshaw (ed) Census Users' Handbook, Longmans, London