Complexity Science and 21st Century Issues
London School of Economics 25-26 March 2004

Dr Bryan D. Gross, President and Chief Executive Officer, MPSI Systems Inc.

"Complexity & Consumer Behaviour"

MPSI Systems Inc is the market leader in retail and consumer science with 10 of the top 20 Fortune 500 companies as clients and more than 400 clients worldwide. Having analyzed over 2 million retail outlets across a wide range of retail operations MPSI is confident in discussing how retailers can meet the needs of their consumers, leading to success in the market place.

It is consumers that drive retail and our assumption has to be that they do it by making logical decisions. If we can test for and respond to those logical needs we can be successful in the market place.

Before we can begin to assess the needs of the consumers, we have to know who they are, what they buy, when they buy, where they buy, how they buy, why they buy, and how much they buy. The ability to answer these questions makes the difference between retail success and failure and since we know that consumers are driven by logical but complex interactions, what we need is a conceptual model that captures that behaviour.

What we use is called a ‘retail consumer value chain’. We chose this for two reasons: one is that it links the vital elements that need to be considered for a successful operation and two, that a chain is only going to be as strong as the weakest link. So if we're looking at consumer behaviour then buying decisions are based on: the location of the retail outlet, the physical characteristics of the facility itself, the price of products, the merchandising or presentation of those products , the operations that surround those products, the brand of those products and the competition that the particular store faces. If any of those things are weak then the whole performance chain is at risk. People often ask the question: 'How much volume of trade would I do in my store if I put it in the right location?' And we often say 'none and infinite' because basically the how well you execute all components of the retail value chain will drive performance more than any single component in isolation. We also know that consumer behaviour is very much dependent on geography and proximity.

There is no panacea for influencing consumer behaviour and when we look at the retail value chain we have to know, market by market, product by product, consumer by consumer, what the needs are because each has a different sensitivity. Modelling particular consumer behaviour and particular needs makes the task very complex.

Retailers need to fit the volume of service or goods to consumer demand. Segmentation and expenditure are very often a function of the way that consumers interact with each other in the population space. A buying fever can spread through a community in very much the same way that a disease spreads. It’s very much a proximal kind of situation; consumers can start behaving like other consumers depending on their proximity to them. So if we look at different kinds of needs across a market place; whether for beer, wine, tobacco, candies, coffee, pastries, fast food, milk and dairy, we can look at the demand in a particular area for those particular products and see how closely a store matches that demand. We know from case histories and experience that the closer the match against the competition the greater the likelihood of retail success. We know that using the concept of the retail value chain in which we look at: location, traffic, facility, merchandising, price frame, operation and so on, that consumers have a specific importance associated with each of those and if we can quantify that then we can give good advice on options.

We are often asked if we can boil down what affects consumer behaviour to the simplest kind of phrase. We know, for example, that the three most important things about a retail store are location, location, location. But that’s not enough. If this were the only consideration then mapping retail store performance against the strength of location would always give a positive correlation between location and performance. If we look at actual data from over 30,000 retail sites in a particular industry segment across the globe and we try to fit a linear model to that data we can see that less than 1% of consumer behaviour is driven by location. Models that can only predict less than 1% of consumer behaviour, are quickly rejected by retailers needing to make sound and accurate network planning decisions.

We are often told that price is the main driver of consumer behaviour and on that basis, we would then expect store performance to degrade as price increases. However, if we plot the data in different retail sectors for 42,000 retail stores across the globe, we do not see a direct correlation when we try to fit a non-linear equation to that. And again, we can barely account for 1% of retail performance by what is considered one of the most fundamental drivers of consumer behaviour. So the simple idea that there are primary drivers of consumer behaviour in a very predictable simple manner is not true.

What we do know is that when we put together a retail value chain with the component parts and link them together in non-linear ways, we can expect a certain kind of behaviour. We expect consumers to respond positively to enhanced location, or facility, or strength of merchandising, or more aggressive pricing and frame structure. When we plot the data, we can see a positive relationship between all of these things put together in the way retail stores are performing. Fitting a global model to the data and taking out the effect of specific regional needs of the consumer, we can account for 30% of the retail store performance.

If we fit a non linear model (i.e., making the data more applicative rather than additive) we can account for nearly 70% of the behaviour in this particular sector and this seems to be perennial and valid across all sectors. Once we do this, we can start predicting to a very high degree of accuracy. By getting more complex, we're predicting better and better the behaviour at the retail sites we're using as experimental units for measuring consumer behaviour. If we then fit a regional non-linear model, where we've allowed the model to calibrate a particular market place, e.g. Chicago, London or Paris, we can accurately predict or account for 81% of the variability in store performance. Up to this point we have dealt with location and proximal effects and a stores ability to meet customer needs. If we now add to this the effect of proximal competition, we can account for 93% of the variability associated with retail store performance. So as the models get more and more complex we can do a better and better job of predicting what consumers are going to do. It bears out the assumption that consumers are very complex creatures. Logical but complex.

Consumers are neither satisfied nor enticed by simple retail solutions and it is important to consider the contribution to performance of each link of the chain of the consumer-value chain. We can see in the example graph, from the presentation, that there is an increasing contribution from location through price, brand, merchandise, operation and facility. It is this kind of retail science that provides the robust objective discipline to unravel consumer complexity issues and as we build the retail value chain we can see that we can do better and better at predicting performance.

When I first entered the applied sector and started dealing with retail analysis I started to utilise some of the very sophisticated classical academic approaches and modelling techniques that I had used before in the academic sector with little success. I won't say that I sold my soul, but I certainly put it out for rent in trying to figure out how to apply academic mathematical concepts to the retail sector and the commercial sector in general. It was also obvious that we had to develop systems, models and analytical solutions that customers needed and could actually use. So we worked on commercial simulation systems for “real” customers rather than just in-house models. It was also important to deliver systems that were deterministic in use. This meant we had to take the very advanced techniques used at the Santa Fe Institute and, with a very large hammer, pound them into something that we could use in the commercial field. This is something we have to bear in mind when we recruit our analysts and modellers from the academic sector. It is often difficult for them to make the transition to the real world. In fact initially it's not a very happy environment for pure statisticians who have been classically trained. But once they start to make the transition and pick up the passion for applying their art or science, they really start to catch on to what it means to fit classical approaches to real world problems.

We use a whole group of different kinds of models to apply science to retail and we know that the volume of data is important in defining the means to ascertain the solution to a retail problem. But our approach if different. Instead of typically using available data and then explaining how it arises, we try to develop a technique that solves a problem and then raid the data base or collect the data. So, in addition to data, we use geographic information systems, statistical methods, spatial interaction models and agent based methods. Geographically, information systems are everywhere now and I don’t think there's anyone in retail or scientific setting that doesn't use them. Statistical methods always provide good solutions and the nice thing about those is that they provide simple solutions. Classical statistics in terms of slivers and segmentation systems and analogising and clustering techniques give rise to spatial interaction variables and we can use simple non-linear techniques to explain things in simple terms to our customers.

Recently we had a question from a retail customer who asked : 'What is the likelihood of hypermarkets invading various markets in the US and to what degree will it impact our retail network?' Our solution was to do a market penetration simulation where we invaded the market hundreds of thousands of times with hypermarkets and we were able to explain exactly how it would bust their existing retail network. The difficulty with the client was that they could explain a simple time series to their management but not the implications of a stochastic simulation. So in a way our techniques are limited by the ability of our clients to understand the significance of the simulation. That is not an ideal situation but it is realistic.

We also use spatial interaction methods and agent-based technologies which offer the best techniques for explaining consumer behaviour in a retail market place. We’ve had a lot of discussion today about the power of agents and agent-based modelling. Basically what we do is start by modelling the individual retail site and the individual consumer or some aggregate set of consumers. By modelling those and allowing them to interact spatially and using the retail value chain, we are able to build a model of the retail system. And the nice thing is we can start with fairly simple equations and see more complex behaviour emerge from the system as we route it through to a higher level. What we found is that agent-based technology in the context of spatial interaction technology allows us both to predict consumer behaviour and also explain it. Those of you who have done progression models and come up with prediction models know it always explains what's happening.

What we've seen from agent-based technology married to spatial interaction is that these models do a very nice job of modelling and explaining the interrelated needs of the consumer, the non-linear responses of the consumers, and the way they self organise. Whether in transient closed systems or segmentation systems the models do a very nice job of explaining that type of behaviour and they explain very well the feedback of retail on consumers and consumers on retail. In this particular setting, it is very much like a genetic situation where the environment can drive the population and the population can modify its environment. Retailers, by changing their offer can impact on consumer behaviour and demand, and as consumers change their demands or needs they then force the retailer to change. So we can see these feedback mechanisms and we can model adaptation. By looking at the way consumers change over time as new things happen and taking into account spatial interaction, we can model the ability of sites to impact on consumers and vice versa.

Agent-based retail modelling provides one of the best techniques for accurately and realistically explaining complex consumer behaviour. One of the examples that I'd like to talk about is what's referred to in retail as the 'brand network' effect. The idea is that, as you build a network of retail stores you can create a synergy of those facilities across time. You might call it a critical mass effect. By building up the number of outlets or outlet share in the market, retailers can actually create non-linear feedback in their performance. Using a static model we only see an aggregation of average facility performance, but using agent-based techniques with feedback mechanisms we can actually simulate the growth of demand and the adaptation and change of consumers across time. That would not be possible if agent-based techniques had not been developed.

The question then, is whether this kind of accuracy and realism can be used for planning and the answer is unfortunately 'no', because in the retail sector there is a problem with the number of ways things can be combined. In trying to realistically forecast consumer behaviour there are so many different ways that consumers can respond and so many different ways a retailer can meet their needs that it becomes a virtually impossible task. Here's an example of that kind of difficulty. We had a multi-national retailer who is very aggressive in choosing locations and building and formatting stores, as well as pricing and marketing. They were looking at a growth market and 150 potential new locations. They had a build quota of six stores per year and a deadline of eight months to put together a three year plan. Using simulation techniques, they were able to create and develop three plans per day assessing all the characteristics of the retail facilities and modifying them to see the outcome. Over the period, they were able to run 500 different plans from which they selected one. Unfortunately, the reality was that there could be 14 billion possible plans to be evaluated, so in the eight months they were only able to look at a very small portion of these. The problem they faced was that every time a plan was presented, someone would say: 'What if Wal-Mart did so and so' and the whole plan would go out the window.

So this is where optimisation comes in. We have a very sophisticated technology called agent-based spatial interaction modelling to assess complex behaviour but alone it doesn’t get us where we need to be in terms of applying a plan. Optimisation enables us to find the best solutions. If we can find a way to marry models and optimisation techniques we can come up with best predictive solutions before going to the market.

There are different optimisation techniques, but when you look at the retail network it’s really like a string of DNA with each store like a gene, and when genetic algorithms are married to the agent-based simulation models we can select and recombine and maximise fitness for millions of evolving combinations. What we find is that the genetic algorithm married to the spatial interaction can dramatically reduce the number of things we have to consider because it looks at ever improving situations.

For example, in 2003 we were contracted to develop a retail plan for a major global retailer in a large US market. The market had over 9 million consumers, 2,000 competitors and 300 stores. What we saw in the market place over the period of time that we studied it, was that they were undergoing an erosion of their position, of market share, total volume as a proportion of other network volumes, and this was forecast to continue.

Salvation strategies were developed by the client. They decided to invest US $300 million and of their 300 company owned and operated retail outlets they chose candidates for site specific strategies. 90 candidates were for demolition and reconstruction, 50 were for re-imaging, 200 were going to be sold, 65 were simply going to close, 193 were protected as franchises and there were going to be 40 new locations. So there were a big number of possible things to consider and a lot of different solutions and there was no way this client could carry out a manual network planning process even with the best of retail simulation technologies. Optimisation was therefore the solution.

Genetic optimisation however enabled them to expedite their planning cycle from eight months to about six weeks which allowed them to consider more possible solutions and significantly reduced the manpower needed. In fact they reduced a team of 25 planners to about 5. Genetic optimisation also increased the objectivity and decreased the subjectivity in finding a solution. The great thing about this method is that not only are the client’s own goals and objectives taken into account, but that these can be very dynamic. The constraints that could drive the solution were very specific to that particular market, that particular sector and for that particular client. Basically what they decided to do was to rationalise their network and energise it by improving the efficiency of their retail stores on the assumption that this would lead to improved profitability.

One of the particular things they wanted to do was retain their total volume, which seemed tough because the idea of rationalising the network by closing stores would seem against it. However using the optimisation method this could be built in and was quite feasible. They also had levels of returns for each individual store and throughput volumes and wanted to find optimal formats for each one in terms of how goods should be priced and stored and so on. Finally they wanted to come in under budget on the analysis which we did.

By taking all their objectives and rules and economic facts and feeding them into the system, genetic algorithms generated site specific strategies which in turn fed the retail model which after many iterations developed an optimum method. In fact, over the course of the simulation, we ran approximately 14 million evolving client store configurations which, with the genetic algorithm, gave an ever improving solution. Over 250 million forecasts were produced not just for the client's own network but for all the competitors' networks so they knew exactly how they would impact on them in terms of spatial interaction. The optimal solution was 27 New-To-Market stores, 32 raze and rebuilds, 16 re-images, 83 closures and 42 rebrandings.

If we look at the relationship of the client's market share to outlet share for their primary product we see it starts in a fairly high position on the curve. In fact their share is near saturation, but optimisation in conjunction with agent-based technology allowed them to rationalise the network so that they could grow their market share further by improving facilities, operations and price relative to the competition. If we look at their other primary product in the same way, they start out in a very poor position because of erosion. By running the operation they were able to minimally improve their outlet share and significantly improve their market share. In fact improvement was made across all categories but product one and two are very high margin categories and that's where they were fairly weak. Given that all they wanted to do for their primary product was maintain their position they even improved that. The same was true of their market effectiveness in that their market share divided by their profit share for each store also significantly improved across the products.

In summarising this case study of optimisation and agent based modelling: the complexity of the retail setting demanded a best-of-practice approach and this involved marrying simulation with optimisation to produce a repeatable network plan which:

- improved network performance
- improved volume
- improved cash flow
- reduced the cost of marketing
- reduced planning and implementation
- reduced capital investment (in fact saved US $100 million)

In general, we can conclude that starting with the assumption that consumers are complex, the need is to simplify. We know that each situation is unique and that there are a number of drivers of consumer behaviour which have to be taken into account. Using agent based modelling and optimisation techniques, we can simplify complex behaviour into quantifiable cause and effect and this 'best of practice' approach provides predictive realism.

<u>Questioner 1:</u> Can we assume that choices are made rationally?

<u>Answer:</u> Well we're all consumers. If we were behaving purely at random then our hypotheses would not enable us to predict consumer behaviour. We might look at the numbers and say initially that as far as store location was concerned, consumers were behaving very randomly. But when we add the other factors and start partitioning up the variability we can actually see for example, that although location is a significant driver, consumers are making their decisions very logically. One of the things that we find is that consumers tend to 'switch on and switch off'. We have an analysis that we do for price sensitivity and we see consumers switching on and off to price as a

primary driver. That means that instead of a nice non-linear continuous curve we start to see discontinuities. But it's still very much the result of logical choice.

Questioner 2:….Inaudible.

Answer : All I can say is that, if we fall back on the analysis that we have done in the last seven years, that if you measure the right things you can predict consumer behaviour.

Comment : All that you mean (by rational) is that people are logical but somehow predictable, but being predictable is different from being logical so it could be any kind of rationality as long as it's predictable.

Answer: So the word you're concerned about is 'logical', but you're happy with 'predictable'.

Questioner 3: My assumption is that it must be hard to justify solutions that are generated automatically and you told us an anecdote about a board member saying: 'What if Wal-Mart had done so and so?' Because you use an automatic optimisation program to generate the solution, I wonder whether you have a rationale to explain why it's a good solution in that situation.

Answer: What we are aiming to do is to track and deliver multiple solutions so that, along with a data base, it enables the client to compare questions in the simulation. We often include competitor reactions, such as the Wal-Mart possibility, and let the optimisation evolve finding the fittest solution.
------------------------------------Loss of sound---

Questioner 4: I think I've got the answer from the presentation, but is there one word which describes what the consumer wants?

Answer: No, and that's a question that's asked of us all the time.

Questioner 4: Well the reason I think that's interesting and informative is that there is a whole body of anthropology where they do similar optimisation models, but take the hunter gatherers, anthropologists always optimise one thing like time energy or food consumption and I've always thought that was wrong.

Answer: Quite honestly this is a similar situation to doing a study on an ecosystem in biology. There is a hunt for resources and consumers feeding off consumers or resources. So there's a very complex interrelated linkage of what we think now is a switching on and switching off of those needs. There is no simple answer to why this happens.

Questioner 5: Perhaps I'm oversimplifying your framework but I gather that in these complex situations that you're able to come up with what are apparently better plans and what are clearly better defensible plans in that they survive the selection. How do you test the accuracy of your plans against the subsequent working out?

Answer: One of the reasons is, that over the last 30 years, we have been able to validate the results of these solutions against actual performance. Time and time again, we have looked at the way the network actually behaves in the wake of prediction and implementation. So basically it's track record.

Questioner 6: How do you control for the different communities of consumers or the heterogeneity of consumers?

Answer: Basically what we look at with each and every homogeneous set of consumers, whether it be a household or a middle age group, is a K factor or function that defines the degree of interaction between one set of consumers and another and between one retail outlet and another and so on.

Questioner 6: How do you model for all this heterogeneity?

Answer: Basically they're flow models, predicting the way that consumers move through a market using distance measurement systems and traffic flows so that we can measure the dynamics.