Toward Automated Pricing and Bundling of Information Goods

Christopher H. Brooks and **Edmund H. Durfee**
Artificial Intelligence Laboratory
University of Michigan
Ann Arbor, MI 48109
{chbrooks,durfee}@umich.edu

Abstract

The introduction of electronic means of sale and dissemination of information goods has made a new set of pricing and bundling strategies possible. In particular, information goods have a negligible marginal cost. This provides producers with an incentive to sell goods in bundles. This new space of strategies also presents producers with a set of problems to solve. In most cases, a producer will need to learn both what prices to charge and what bundles of goods to offer. We discuss how a producer can balance the complexity of a price schedule against the time needed to learn it, with the goal of maximizing its aggregate profit. We then examine how a producer can use taxonomic information about the content of information goods to help discover which goods should be offered for sale and how they should be bundled together. We discuss results to date and present a set of questions for further research.

Introduction

The growth of the Internet as a medium for conducting commerce has presented researchers with a new and challenging set of problems involving how to conduct electronic commerce and determining which aspects of electronic commerce can be automated. The incorporation of autonomous agents into electronic commerce provides an opportunity for more complex and dynamic pricing and negotiation schemes, on-the-fly repackaging and bundling of related products to match consumer preferences, and the targeting and development of specialized niche markets which would not normally be feasible in a physical environment. Of course, all these new opportunities are accompanied by a host of related technical problems which must be overcome if electronic commerce is to be successful.

We are currently studying these problems as they relate to information economies. More particularly, we are interested in how a producer of information goods (such as movies, news articles, or web queries) can attract the best set of consumers, where "best" are those consumers which maximize the producer's long-term profits. In order to attract consumers, a producer must make decisions regarding which goods it should offer, how they should be priced, and how they should be bundled together. These decisions are based on the producer's beliefs about consumer preferences (and its knowledge of those preferences), its beliefs about the strategies that other producers will take, and the information to be gained from a particular offering.

Since a producer typically does not know either the preferences of consumers or the strategies of other producers, this knowledge will need to be inferred. Furthermore, this process of inferring preferences and strategies and then using this knowledge to determine what to offer and how to price it will take place within a dynamic environment. Consumers will enter and leave and competing producers will change their strategies. This leads to a moving target learning problem (Vidal & Durfee 1998b), where the function an agent is trying to learn changes over time. In this case, a producer will want to perform well *while* it is learning, since the target function may change before the producer has time to learn it fully.

As an example, consider a scenario with two providers of cable TV channels and a population of heterogeneous consumers. Each provider will choose a set of channels to offer for sale, from generalized packages which include many different types of programming to specialized packages targeting particular markets, such as a package of sports channels or cooking channels. Consumers will then choose which packages to purchase, based on their valuation of the content of these channels. Consumers may even drop out altogether if there are no satisfactory offerings. In making a decision as to what to offer, a cable TV provider must consider three factors: first, what does it believe the consumer population values? It makes no sense to offer channels no one will watch. Second, what are its competitors offering? It also makes no sense to offer the same packages as your competitors, but at a higher price. There may be instances in which the best strategy is to compete directly with a competitor for a particular niche, but it is likely that in many other instances the best strategy is to target a niche which is not being satisfied by other producers' offerings. Third, what will a provider learn from offering a particular bundle of channels? Since a provider typically does not know exactly what information content is preferred by consumers, it must explore the space of possible offerings. Even if a particular bundle of channels does not sell particularly well, it may be very useful in helping the provider to learn what consumers do (and do not) prefer.

Traditionally, these decisions have been made by humans. Cable TV providers change their programming very infrequently, since there is a large cost to doing so. The incorporation of autonomous agents into the electronic packaging and distribution of information changes this dynamic, making it possible for providers to reorganize their offerings more easily, while hiding this complexity from the humans in the system, be they producers or consumers.

In this paper, we begin by discussing how an information goods producer can select a pricing strategy to use, balancing the profits to be gained from learning a particular schedule against the complexity of learning that schedule. Following this, we discuss some characteristics of information goods that producers can exploit when learning about consumer preferences over these goods. We describe how a producer can use taxonomic information about information goods to potentially speed up their learning, and describe our current research regarding learning of consumer preferences in competition with other producers.

Information Goods

Information goods, when distributed electronically, are typically assumed to have a high fixed, or first-copy, cost, and a negligible marginal cost. This means that once one copy is purchased by a producer, additional copies can be created virtually for free. This makes *bundling* of information goods an appealing strategy; if a producer can reproduce articles at a very small cost, it can create a bundle of different articles which appeal to different sets of consumer tastes. For example, rather than selling a sports article and a news article separately for $1 each, a producer could sell both together for $1.50. A consumer who valued the sports article at $1 and the news article at $0.50 (or vice versa) would buy the bundle, yielding more profit for the producer than if the articles were sold individually at $1 apiece, where the consumer would only buy one article. This opens up a whole new set of potential pricing schedules for a producer. The first question we have examined is: how a producer can select from among these schedules, given that it does not typically know consumer valuations?

Learning Price Schedules

Initially, we examined a system in which an information goods producer offered a set of goods according to a fixed price schedule and consumers were able to choose the subbundle of goods that they wanted from this set. This freed the producer from the problem of selecting which goods to offer and allowed us to focus on the question of determining what price schedule a producer should employ.

We considered five pricing schedules:

- Per-article (linear) pricing. Consumers pay a fixed price for each article.
- Pure Bundling. Consumers pay a fixed price for the entire collection of articles.
- Two-part tariff. Consumers pay an entry fee, plus a per-article price for each article.
- Mixed bundling. Consumers are offered a choice between a per-article price and a pure bundle price.
- Nonlinear pricing. Consumers pay a different price for each article consumed.

The first two schedules have one free parameter, the second two have two free parameters, and the last has N free parameters, where N is the total number of articles for sale. In a steady-state, perfect information system with heterogeneous consumer tastes, price schedules with more free parameters allow a producer to extract more profit, since they allow a producer to more accurately fit its price schedule to the consumer demand curve. (See (Brooks *et al.* 1999) for details). Since a producer has perfect information, the problem reduces to finding the schedule with the highest equilibrium profit and always charging that at each iteration.

When the producer has to learn the parameters of a price schedule, the problem changes. In this case, the producer must concern itself with aggregate profit. It is not enough to eventually find a profitable schedule; a producer should try to maximize its *total* revenue over time.

We conducted a set of experiments to determine both how difficult each price schedule was to learn and also how much aggregate profit a producer would accumulate while learning. We used two different learning methods: a neural network and amoeba (Press & others 1992), a hill-climbing algorithm for nonlinear optimization. A significant difference between these algorithms is that the neural network tries to learn the entire profit landscape (mappings from prices to profit for all inputs), whereas amoeba strictly tries to find the global optimum of the profit function without retaining any state about the landscape. For problems with a static consumer population, the exploration performed by the neural network turns out to be unnecessary; optimization is all that is needed. However, if the consumer population (and therefore the profit landscape) is changing, then it can be advantageous to retain information about its structure. Results for each algorithm learning each price schedule are shown in Figures 1 and 2.

While the two graphs have some differences, the high-level results are the same. If the producer has more than a few iterations to interact with the consumer population, the two-parameter schedules significantly outperform the one-parameter schedules. Also, nonlinear pricing (requiring 10 parameters in this instance) is quite difficult to learn;

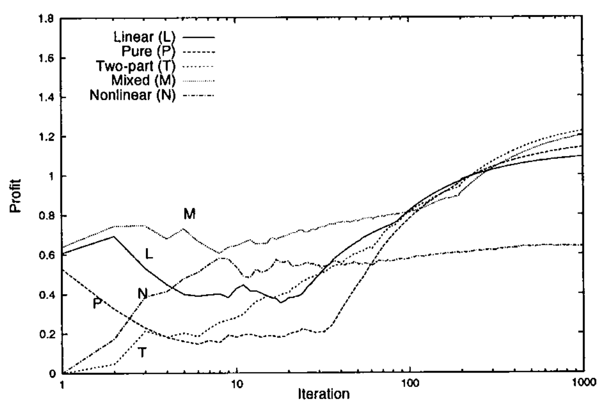

Figure 1: Aggregate Profit per Iteration for the Neural Net

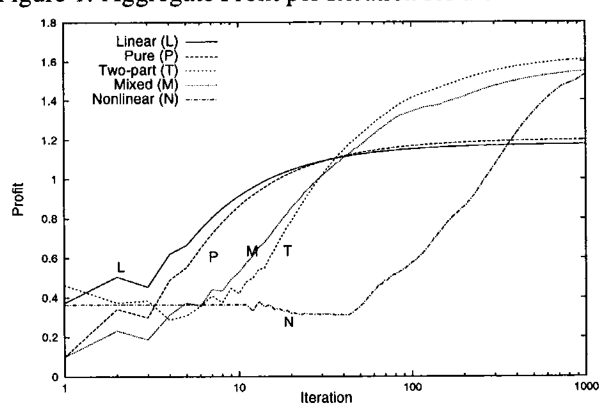

Figure 2: Aggregate Profit per Iteration for Amoeba

its superior steady-state profits never make up for the time needed to learn them. Also, two-part tariff outperforms mixed bundling, even though the two schedules have identical steady-state profits. The reason for this is that the two parameters are more tightly coupled with two-part tariff; if a producer raises the subscription price slightly, it must lower the per-article price to compensate. With mixed bundling, the producer is offering a choice of two one-parameter schedules, and so a small change in the per-article price may not necessarily yield any information about consumer preferences over bundle prices.

In this problem, we assumed that the producer has access to a large store of articles which it can offer for sale. Also, we assumed that consumers had the ability to sort through this store of articles and find the ones they want. This allowed us to focus on the problem of learning prices. However, these assumptions may not be realistic. Typically, a producer will need to acquire articles from an outside content provider. More significantly, consumers do not typically have the ability to costlessly search through a large set of articles to find the set they are interested in. Instead, they will have to pay a processing cost, either in terms of time or storage, which will decrease the value a consumer ascribes to the set of articles as a whole. If the producer could offer a subset of articles which were highly valued, it could reduce the consumer's processing cost and therefore extract more profit. This leads the producer to consider the problem of which goods to offer, and how to offer them. We conjecture that producers of information goods can bundle together goods with correlated values to deal with this.

Bundling According to Content

Previous research (Fay & MacKie-Mason 1998) (Bakos & Brynjolfsson to appear) has shown that, under certain assumptions about consumer heterogeneity, if consumers are able to freely dispose of unwanted articles, pure bundling (selling all available articles) will always yield at least as much profit as per-article pricing. Empirical work done on the PEAK project (Riveros 1999) (MacKie-Mason, Riveros, & Gazalle 2000), provides evidence that, even when consumers must pay a disposal cost, *sub-bundling*, or offering a subset of available articles as a bundle, can be an effective strategy. This should come as no surprise; we see this strategy employed every day by academic journals, specialty cable channels such as ESPN, and websites such as Yahoo and About.com, which group together links to websites with related content. This opportunity for sub-bundling, particularly when it is being performed automatically by agents, leads to a problem which has not been examined within the bundling literature: how does a producer determine which goods to bundle together?

One aspect of information goods which we exploit in tackling this problem is the fact that they are primarily valued according to their content. We assume that consumers have a set of *information needs* (Salton & McGill 1983) which can be satisfied by goods with a particular sort of content. For example, a consumer might want a description of the Quicksort algorithm, or a review of currently playing movies, or news about Spring Training. A producer can then construct categories of articles that contain particular content and use this as an aid in bundling. For example, our cable TV provider might find it useful to group the Olympics, pro basketball, and pro football into a single category labeled "Sports" and sell these programs as a bundle.

Producers can also exploit the relationship between categories when constructing bundles. By creating sub- and super-categories, an information goods producer can construct a taxonomy of different categories of articles, as shown in Figure 3. This ability to then specialize or generalize a bundle's content allows an information goods producer to either specialize in a particular niche market or to generalize and capture consumers with heterogeneous preferences. Our cable TV provider might determine that customers value pro basketball much more than football and decide to offer a separate "NBA" bundle containing only NBA games, which could then be sold at a higher price than the more general "Sports" bundle. On the other hand, if the consumer population contained a set of users with interests in both basketball and movies, the producer could construct a more general "Entertainment" bundle, containing a more diverse collection of articles, which would appeal to these consumers.

Note that the decision to bundle related goods does not free the producer from deciding upon a price schedule. The producer now must decide how to price these different sub-bundles. For example, it might choose to use two-part tariff, charging an entry fee plus a flat rate for each sub-bundle, or mixed bundling, with a fixed rate for each sub-bundle or a separate price for access to all sub-bundles. All this issues discussed previously still apply; the distinction is that the price is being applied to bundles of articles rather than individual articles.

One question we might ask is: if an information goods producer can create all these sub-bundles as needed, why doesn't it just offer all possible sub-bundles and avoid doing any learning? There are three reasons for this. First, this would result in a nonlinear price schedule, where each sub-bundle (potentially) had a different price. As we discussed above, this would impose a serious burden on the consumer; he or she would be forced to look through the power set of all articles to determine which package or packages to buy. Even with the aid of automation, this seems like an onerous task. Second, if the information goods producer is not also able to engage in price discrimination (charging each consumer a different price), it may be in the producer's best interest to not offer certain bundles. An advantage of bundling is that it helps to smooth out heterogeneity in consumer preferences. If all possible sub-bundles are offered, then this advantage is lost. Third, recall that an information goods producer must still determine how to price all of these bundles. As we saw previously, the more parameters a pricing schedule has, the more difficult it is to learn. Exactly how many sub-bundles a producer should offer to make the problem "easy enough" while still extracting a significant fraction of the available profits is still an open question.

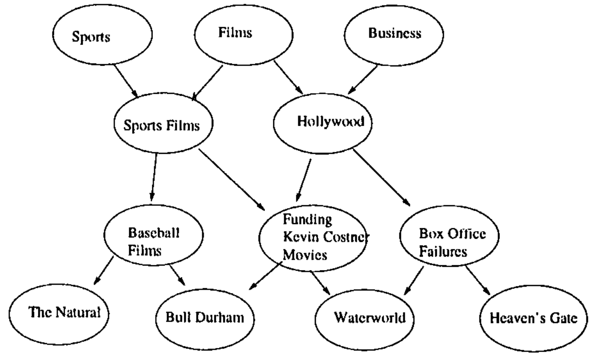

Figure 3: A Taxonomy of Information Goods

A premise of our research is this: consumer information needs correspond to categories of articles. Therefore, a producer which learns which categories of goods to offer is learning consumer preferences for those goods. Further, if a producer knows the taxonomic relationship between categories, it can exploit this knowledge to improve its decision as to which bundles of articles to offer. In the simplest case, a producer can use a taxonomic structure as a guide in searching the space. In a more complicated scenario, the taxonomy could be used to provide additional evidence for a producer that wants to engage in belief revision regarding its model of consumer preferences.

This problem is potentially extremely complex. The consumer population typically has multiple, heterogeneous information needs. Consumers may not use the same taxonomic relationships that a producer uses. Also, other producers may also be trying to satisfy these needs. A producer will want to quickly determine a sub-bundle of goods which, given the strategies of other producers, yields it as much aggregate profit as possible. We conjecture that, in many cases, producers will be better off specializing and offering sub-bundles which appeal to particular niches of the population, rather than competing as generalists for the population as a whole. Each producer's problem can then be viewed as one of identifying niche markets.

Learning Consumer Preferences using Taxonomic Information

We are taking a decision-theoretic approach to the question of how an information goods producer can learn the correct set of bundles to offer (and how to price them). As in the previous experiments on learning prices, the producer is interested in maximizing its aggregate profit. Therefore, it wants to select a series of bundles which both lead it to an optimal solution and provide satisfactory performance in

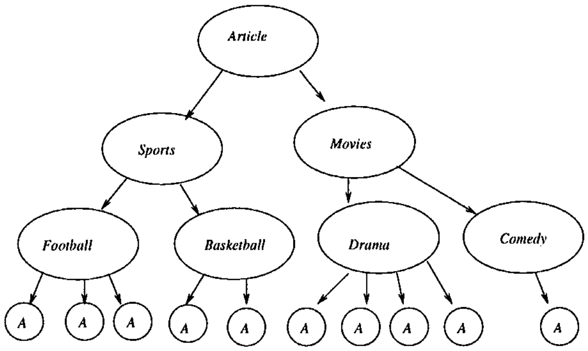

Figure 4: A simple taxonomy of article categories.

the transition. A producer can use the taxonomic information it has about the relationship between article categories to help make this decision and update its beliefs about consumer preferences. By postulating the possible effects of offering each category, a producer can estimate both the expected utility of offering that category as well as the value of the information which is gained from offering that category. This information is valuable to the extent that is helps the producer make better choices as to what category to offer in future iterations.

An example may help to make this clearer. Consider the taxonomy in Figure 4. This is about the simplest possible taxonomy: a binary tree. Let us also assume that there is a single producer and a single consumer. Initially, the producer has no knowledge about consumer preferences, and so all categories seem equally likely to match the consumer's information need. Finally, to make the example simple, let us assume that articles are dispensed to the producer at a rate of 3:2:4:1 for football, basketball, drama and comedy, respectively. It can then choose subsets of these to bundle and offer to the consumer.

If the producer only gets one chance to offer a bundle to the consumer, it can simply calculate the expected utility for each category. For example, (assuming consumer valuation is 1 for articles meeting its information need and 0 otherwise), if the producer offers Sports, the consumer will be completely satisfied if its need is for Sports, or Articles. If the consumer's need is for basketball, $\frac{2}{5}$ of the articles will be of value, and if the consumer's need is for football, $\frac{3}{5}$ of the articles will be of value. If the consumer's need is for Movies, Drama, or Comedy, none of the articles will have value. This leads to $EU(Sports) = \frac{3}{7}$. Unfortunately for our producer, every category has the same expected utility, so in a one-shot game, the producer will do just as well flipping a coin as reasoning about the consumer.

If the producer will have multiple interactions, this changes, since it can use the results of previous interactions to alter its choice set. For example, if the producer offers Sports and the consumer's need is for baseball or football, the producer will know this with certainty and can offer that in the next iteration. If the consumer's need is for Sports or Articles, the producer has found a successful bundle. If the consumer's need is for Movies, Comedy, or Drama, the producer knows it is one of these, but not which one.

Therefore, when the producer makes its initial decision as to what to offer, it can consider the value of the information it will gain as a result of this first experiment and how this knowledge will improve its second experiment. In a two-period game, for this example, Sports, Article and Movies all yield an expected two-period utility of $\frac{9}{7}$, whereas the leaf categories yield an expected two-period utility of $\frac{8}{7}$.

This example is extremely simple, but it should serve to illustrate how a producer can use this taxonomic information to help infer consumer preferences. Particularly important is the exploiting of the sub- and superclass relationship. If a consumer likes an article of category c, it will like articles from a subcategory of c. Similarly, if a consumer does not like articles of category c, offering a supercategory provides a principled way of searching through this taxonomy.

Open Questions

As this research is still in the early stages, we are presented with more questions than answers. Even the simple preceding example provides us with questions.

We ignored the learning of prices in this example. A producer will typically need to determine both a bundle or bundles to offer and a price schedule. Should it learn the consumer demand a particular bundle before offering other bundles, or learn what to charge and which bundles to offer simultaneously? Also, are there mechanisms akin to the Vickrey auction (Vickrey 1961) which can induce a consumer to reveal its true valuation for a bundle and sidestep the process of learning prices?

We also assumed a single consumer in the previous example. Heterogeneous consumers make the problem more difficult; the producer will receive noisier feedback in this case. In particular, it is not obvious when the correct bundle is found. The producer must instead make a probabilistic determination as to whether to explore further in the space of bundles, based upon the profit received and their estimates of the maximal potential profit for a bundle.

Also, as discussed previously, in a real-world information economy, a producer might be competing with other information producers for the consumer population. Therefore, a producer's calculation of the expected utility of offering a bundle of articles from a particular category will be contingent upon the strategies of other producers. This means that a producer will need to reason about the strategies and knowledge of other producers, which leads us to the question of how much modeling of other producers is enough, or, alternatively, when does further modeling of other producers fail to help, or even harm a producer's progress? (Vidal & Durfee 1998a) provides hope that a producer may be able to shallowly model another producer and profit, while

(Wellman & Hu 1998) warns that agents which try to learn models of each other could wind up being worse off than if they had not learned at all. Clearly, this is a situation to be avoided if possible.

The previous example also assumed that the producer and consumer shared the same taxonomy of articles. This seems unlikely; in fact, it seems more reasonable to assume that different consumers might have different ontologies. A consumer who is an NBA fan will have a very detailed taxonomy of basketball articles compared to someone who is not interested in pro basketball. Most likely, a producer will begin with a taxonomic model which is roughly correct, meaning that it corresponds somewhat to the aggregate consumer model, but will need to be pruned and refined. This leads us back to the question of when it is useful to model another agent. If the producer's model is completely wrong, is it better off not using the model?

Finally, the previous example showed the producer offering a single category. When the producer is interacting with multiple, heterogeneous consumers, it may be able to extract more profit by offering multiple bundles. This actually creates an interesting problem: by offering more different bundles, the producer may reap more short-term profit, but it will potentially learn less about consumer preferences. What is the optimal number of bundles to offer at once? Should this change as the producer's knowledge about the consumer population changes?

Conclusion

In this paper, we have described our ideas and approach to solving the problem of how a producer can learn which prices and bundles to offer in an information economy. We have provided some evidence that the solution which appears best in equilibrium is not necessarily the best; instead, a producer must consider the cost of learning and the amount of interaction it is likely to have with a population when determining how to model a consumer population.

We have also described our plan for allowing a producer to use the content of information goods to make inferences about their value. While this work is still in progress, we believe that our approach will be fruitful, and have outlined questions we plan to address.

Acknowledgments

This work was supported in part by an IBM University Partnership Grant and by the National Science Foundation under grant IIS-9872057.

References

Bakos, Y., and Brynjolfsson, E. to appear. Bundling information goods: Pricing, profits and efficiency. *Manufacturing Science.* also avaliable at: http://www.stern.nyu.edu/ bakos/big.pdf.

Brooks, C. H.; Fay, S.; Das, R.; MacKie-Mason, J. K.; Kephart, J. O.; and Durfee, E. H. 1999. Automated strategy searches in an electronic goods market: Learning and complex price schedules. In *Proceedings of ACM EC-99.*

Fay, S. A., and MacKie-Mason, J. 1998. Competition between firms that bundle information goods. Unpublished.

MacKie-Mason, J. K.; Riveros, J. F.; and Gazalle, R. S. 2000. Pricing and bundling electronic information goods: Experimental evidence. In Compaine, B., and Vogelsang, I., eds., *Internet and Telecommunications Policy Research.* MIT Press.

Press, W. H., et al. 1992. *Numerical Recipes.* Cambridge University Press.

Riveros, J. F. 1999. Bundling and access to information goods: Consumer heterogeneity and marginal costs. Preliminary draft.

Salton, G., and McGill, M. J. 1983. *An Introduction to Modern Information Retrieval.* New York: McGraw-Hill, Inc.

Vickrey, W. 1961. Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance* 16:8–37.

Vidal, J. M., and Durfee, E. H. 1998a. Learning nested agent models in an information economy. *Journal of Experimental and Theoretical Artificial Intelligence* 10.

Vidal, J. M., and Durfee, E. H. 1998b. The moving target function problem in multi-agent learning. In *Proceedings of the Third Annual Conference on Multi-Agent Systems.*

Wellman, M., and Hu, J. 1998. Conjectural equilibrium in multiagent learning. *Machine Learning* 33.