

Two-sided learning in an agent economy for information bundles

Jeffrey O. Kephart¹, Rajarshi Das¹, and Jeffrey K. MacKie-Mason²

¹ IBM Research

PO Box 704, Yorktown Heights, NY 10598

{kephart,rajarshi}@watson.ibm.com

² School of Information and Department of Economics

University of Michigan, Ann Arbor, MI 48109

jmm@umich.edu

Abstract. Commerce in information goods is one of the earliest emerging applications for intelligent agents in commerce. However, the fundamental characteristics of information goods mean that they can and likely will be offered in widely varying configurations. Participating agents will need to deal with uncertainty about both prices and location in multi-dimensional product space. Thus, studying the behavior of learning agents is central to understanding and designing for agent-based information economies. Since uncertainty will exist on both sides of transactions, and interactions between learning agents that are negotiating and transacting with other learning agents may lead to unexpected dynamics, it is important to study two-sided learning.

We present a simple but powerful model of an information bundling economy with a single producer and multiple consumer agents. We explore the pricing and purchasing behavior of these agents when articles can be bundled. In this initial exploration, we study the dynamics of this economy when consumer agents are uninformed about the distribution of article values. We discover that a reasonable albeit naive consumer learning strategy can lead to disastrous market behavior. We find a simple explanation for this market failure, and develop a simple improvement to the producer agent's strategy that largely ameliorates the problem. But in the process we learn an important lesson: dynamic market interactions when there is substantial uncertainty can lead to pathological outcomes if agents are designed with "reasonable" but not sufficiently adaptive strategies. Thus, in programmed agent environments it may be essential to dramatically increase our understanding of adaptivity and learning if we want to obtain good aggregate outcomes.

Keywords: Information economy, information bundling, two-sided learning, economic agents.

1 Introduction

Extraordinary rates of advance in computation and communication technologies¹ have fueled the much-noted proliferation of electronic commerce. Within a few years, we anticipate that software agents will participate in a wide variety of commercial transactions, and may even become economic players in their own right [5]. One important domain for agent economies is the production and distribution of information goods and services, such as news articles, entertainment and other service reviews, and instructional materials. Digital information goods are unusually configurable, and negotiations over the composition and prices for bundles are a natural application for software agents. However, to do so will require that broker agents be able to explore not just prices, but also locations in a multi-dimensional bundled product space.

Very little is understood about strategic search over the joint information good price and product space (that is, search for which products to offer, in which combinations, at what prices).² Agents competing in this space must learn about the distribution of preferences across a heterogeneous and changing customer population, must learn about the strategies being followed by their competitors, and then must optimize their strategies to take into account the new understanding of customers and competitors. Consumers, on the other hand, must learn about the quantity, quality and price of bundled items offered by the various producers. Learning and optimization are both difficult search problems, and in the market context they are closely related.

In earlier work we considered strategic pricing for bundles when all of the relevant parameters are known by brokers [3]. We distinguished between conditions in which firms prefer to offer comprehensive bundles at a single price and those in which they prefer to sell items individually. Although ours is the first work focusing on competitive strategies, the bundling and pricing of information goods has been an active area in recent research [1], [2]. We have also studied price dynamics when competing brokers search a restricted price and product space [5, 6, 4]. Price instability, characterized as “price wars”, was prevalent.

In this paper we examine a considerably more general problem, although for this initial foray we maintain some simplifying restrictions. We envision an agent economy with one information broker agent and many consumer agents. The consumer agents are heterogeneous: they value each item differently, and these values are drawn from a different distribution for each agent. To set profit-maximizing prices, the broker agent desires, but may not know, the parameters of the consumers’ valuation distributions. The consumers also may not know their valuation parameters until they gain experience with the broker’s offered information goods. Therefore, both sides wish to learn. The broker can set prices strategically, learning about consumer valuations from their purchasing behavior at different prices. The consumers can purchase strategically, learning about the distribution of information good values by sampling.

¹ See, e.g., [7].

² For some recent work on multi-agent search, see, e.g., [14, 13].

This environment is very rich and will permit us to explore many interesting questions. In this paper we focus on one surprising result: that when consumer agents follow a plausible but overly naive learning strategy, even if the producer is fully informed (but also somewhat naive), the economy can continuously degenerate with disastrous overall performance. We find a simple explanation for the initial failure, and develop a simple improvement to the producer agent’s strategy that largely ameliorates the problem. But in the process we learn an important lesson: dynamic market interactions when there is substantial uncertainty can lead to pathological outcomes if agents are designed with “reasonable” but not sufficiently adaptive strategies. Thus, in programmed agent environments it may be essential to dramatically increase our understanding of adaptivity and learning if we want to obtain good aggregate outcomes.

This paper is the first in a series of studies directed towards understanding and developing robust adaptive learning techniques that are effective for individual agents and lead to acceptable collective (market) behavior. We are also extending the model to study the much more common setting in which there is competition among multiple producers of information goods. Of course, this setting exacerbates the learning problem, since the producers need to learn about each other’s strategies in addition to learning about consumer valuations.

In the next section we discuss the model. We introduce two relevant distributions: first, a parameterized distribution g from which a given consumer’s valuations are drawn, and second, a distribution h describing the population of consumers — the distribution from which an individual consumer’s valuation distribution parameters are drawn. For tractability, we assume that the broker is restricted to two-part tariff pricing schemes [9]: it can charge consumer agents a subscription price to examine the current items, and a per item price for each item the consumer subsequently purchases.

In a general analysis section, we derive the optimal strategies for fully informed broker and consumer agents as functions of the article value and consumer type distributions g and h . Then in Section 4 we explore specific cases of exploration and exploitation through analysis and simulation. When agents are underinformed and engage in learning, we consider plausible (but not necessarily fully optimal) agent strategies, and examine the resulting system outcomes. We close with a summary of our current results and plans for future work.

2 Model

A single producer periodically (at discrete times $t = 0, 1, 2, \dots$) generates sets of N articles. It sets a subscription fee F and a price schedule $\mathbf{P} = \{P(1), \dots, P(N)\}$, where $P(k)$ represents the price it charges for a subset of $k \in \{1, \dots, N\}$ of the N articles.

At a given time t , each of M consumers are informed about (F, \mathbf{P}) , and they decide whether to subscribe. Then, each subscribing consumer receives abstracts of all N articles, and uses them to assess its value w_j from reading each article, for $j \in \{1, \dots, N\}$. We assume that these values are generated randomly ac-

according to a distribution $g(\mathbf{q}_i; w)$, where \mathbf{q}_i is the set of parameters that define the distribution for consumer i . The parameters \mathbf{q}_i that represent consumer i 's valuation distribution are themselves generated randomly prior to time 0 from a distribution $h(\mathbf{r}, \mathbf{q})$. Once the \mathbf{q}_i parameters are generated for consumer i , they remain fixed for the rest of time (though the agents may not know their true values for a while, if ever).

After assessing the value of each article, a subscriber decides which articles to purchase. It does this by choosing a set of articles K to maximize surplus $s = (\sum_{j \in K} w_j) - P(|K|)$. Henceforth we assume that the articles have been sorted by consumer j so that the w_j are ordered highest to lowest, and thus the set K consists of the first $k = |K|$ articles.

The subscription decision depends on consumer *expectations*. If a consumer believes correctly that its valuations are drawn from a distribution with parameters \mathbf{q}_i , then its expected surplus from purchasing k articles would be

$$\langle s_k \rangle = \left\langle \sum_{j=1}^k w_j \right\rangle - P(k). \quad (1)$$

The consumer can then derive the vector of values p_k , the probabilities that any k is the expected surplus maximizing number of articles. Then the consumer's optimized expected surplus is

$$\langle s \rangle = \sum_{k=1}^N p_k \langle s_k \rangle. \quad (2)$$

The consumer should subscribe if and only if the expectation $\langle s \rangle$ exceeds the subscription fee F .³

If consumer i does not know its own \mathbf{q}_i for the articles offered by this broker, then values \hat{w}_j are the consumer's *beliefs* drawn from a distribution $g(\hat{\mathbf{q}}_i; w)$ after reviewing the abstracts, where the $\hat{\mathbf{q}}_i$ are the agent's current best estimate of the valuation parameters \mathbf{q}_i . When the agent purchases and reads articles, it learns their true values w_j and can then use this sample information to update its beliefs $\hat{\mathbf{q}}_i$ about the distribution of article values. Therefore, a good consumer strategy should take into account the value of learning. For example, when uncertainty about \mathbf{q}_i is high, the consumer might deliberately subscribe even when its estimated surplus is less than F , simply to experience more articles to improve its valuation estimates. Or, having subscribed the consumer might purchase more articles than would maximize expected surplus from current reading.

Turning to the producer's problem, it can choose a subscription fee F and a price schedule \mathbf{P} in each period. These should be chosen to maximize some

³ We assume consumers are *risk neutral*. A *risk averse* agent would want to optimize the expectation of a concave function of surplus, and optimal behavior would depend on second and possibly higher moments of the induced distribution of surplus, not just the expected surplus. See, e.g., [8].

function of expected current and future profits, where current profit can be expressed as:

$$\pi = \sum_{i=1}^M \theta_i (F + P(k_i) - C(k_i)). \quad (3)$$

with $\theta_i = 1$ if consumer i chooses to subscribe, and 0 otherwise. The cost of delivering k_i articles to consumer i is denoted as $C(k_i)$.

In performing its maximization, the producer must take into account the effect of \mathbf{P} and F on the consumer's subscription decisions, and the effect of \mathbf{P} on the distribution of k_i across the set of subscribers. Higher prices will decrease the number of subscribers and the expected number of articles purchased, of course. To compute the optimal subscription fee and price schedule the broker wants to know the distribution $h(\mathbf{r}, \mathbf{q})$ from which the consumers' parameters were generated, and the consumer strategies for subscribing and purchasing. Based on its current beliefs about \mathbf{r} and consumer strategies, the broker can simulate a consumer population and its responses to various (F, \mathbf{P}) schedules, and then pick (F, \mathbf{P}) to maximize a value function. In practice, the broker may not know the consumer type parameters \mathbf{r} , nor the consumer strategies. Thus, the broker may choose (F, \mathbf{P}) to balance current expected profit according to equation (3), against the increase in expected future profit from learning about consumer preferences by observing their behavior when confronted with varying (F, \mathbf{P}) combinations.

3 General Analysis

In this section, we first analyze the expected surplus and number of articles purchased per subscription period for a rational, fully-informed consumer with a given valuation distribution g . Then, we derive an expression for a monopolistic producer's expected profit as a function of its price schedule and the distribution h of valuation parameters \mathbf{q} across the consumer population, assuming that all consumers are rational and fully-informed about their valuation distribution. A rational, fully-informed producer would choose its price schedule so as to maximize its expected profit.

Suppose that a given consumer has its valuations w drawn from a probability density function $g(\mathbf{q}; w)$. (The distribution parameters \mathbf{q} may vary from one consumer to another.) For the sake of simplicity, we assume that the producer constrains itself to a linear price schedule: $P(k) = k\Delta$. Then a rational consumer will purchase k articles, where k is the number of articles with valuations exceeding the threshold $P(k) - P(k-1) = \Delta$. The probability p_k for exactly k articles to have valuations $w > \Delta$ is

$$p_k = \binom{N}{k} G(\mathbf{q}; \Delta)^{N-k} (1 - G(\mathbf{q}; \Delta))^k \quad (4)$$

where $G(\mathbf{q}; w)$ represents the cumulative distribution function that corresponds to $g(\mathbf{q}; w)$, i.e. it is the probability for an article to be valued at less than w .

From this we can compute the expected number of articles purchased:

$$\langle k(\mathbf{q}, \Delta) \rangle = \sum_{k=1}^N k p_k = N [1 - G(\mathbf{q}; \Delta)] \quad (5)$$

where the last equality follows from simple manipulations of binomial coefficients.

The expected surplus $\langle s \rangle$ can be obtained from Eq. 2, provided that we first compute $\langle s_k \rangle$, the expected surplus given that exactly k articles prove to have valuations $w > \Delta$. The conditional probability distribution for a single draw from $g(\mathbf{q}; w)$ given that $w > \Delta$ is

$$\tilde{g}(\mathbf{q}; w) = \frac{g(\mathbf{q}; w)}{1 - G(\mathbf{q}; \Delta)} \Theta(w - \Delta) \quad (6)$$

where $\Theta(x)$ represents the step function, equal to 1 if $x > 0$ and 0 otherwise. The expected valuation for this distribution is

$$\bar{w}(\mathbf{q}; \Delta) = \int_{\Delta}^{\infty} dw w \tilde{g}(\mathbf{q}; w) \quad (7)$$

The expected sum of k draws from this conditional distribution is $k\bar{w}(\mathbf{q}; \Delta)$. Subtracting the price $P(k) = k\Delta$, we obtain:

$$\langle s_k \rangle = k(\bar{w}(\Delta) - \Delta) \quad (8)$$

Inserting this result into Eq. 2, we obtain:

$$\begin{aligned} \langle s(\mathbf{q}, \Delta) \rangle &= \sum_{k=1}^N p_k k [\bar{w}(\mathbf{q}; \Delta) - \Delta] = [\bar{w}(\mathbf{q}; \Delta) - \Delta] \langle k(\mathbf{q}, \Delta) \rangle \\ &= N [\bar{w}(\mathbf{q}; \Delta) - \Delta] [1 - G(\mathbf{q}; \Delta)] \end{aligned} \quad (9)$$

Note that the consumer will subscribe if and only if $\langle s(\mathbf{q}, \Delta) \rangle$ is greater than the subscription fee F .

Now we take the producer's perspective. If we assume (for simplicity) that the cost of producing and delivering k items is $C(k) = k\gamma$, then the expected profit is

$$\begin{aligned} \langle \Pi(F, \Delta) \rangle &= \sum_{i=1}^M \Theta(\langle s(\mathbf{q}_i, \Delta) \rangle - F) (F + (\Delta - \gamma) \langle k(\mathbf{q}_i, \Delta) \rangle) \\ &\approx \int_{\mathbf{q}} d\mathbf{q} h(\mathbf{r}, \mathbf{q}) \Theta(\langle s(\mathbf{q}, \Delta) \rangle - F) (F + (\Delta - \gamma) \langle k(\mathbf{q}, \Delta) \rangle) \end{aligned} \quad (10)$$

where $h(\mathbf{r}, \mathbf{q})$ is the probability distribution for the consumers' \mathbf{q} parameters, as defined previously. In effect, the last approximation replaces the actual realized set of consumer distribution parameters $\{\mathbf{q}_i\}$ with an ensemble average over all

possible realizations of a set of M consumers generated from the distribution h , and this approximation grows increasingly accurate in the limit of large M . Note that a producer that knows h can compute this profit landscape. A fully knowledgeable and rational producer would set Δ and F so as to maximize $\langle \Pi(F, \Delta) \rangle$.

4 Rational and Bounded-Rational Players

In the remainder of the paper, we shall assume simple functional forms for g and h , and explore what happens when the consumers and/or the producer must learn these distributions. In particular, we suppose that g is a one-parameter exponential distribution given by $g(\mu; w) = \mu e^{-\mu w}$, and that $h(\mu)$ is a uniform distribution in the interval from μ_{\min} to μ_{\max} .

4.1 Fully-informed producer and consumers

As a reference point, we analyze the case where the consumers are fully informed about their individual values of μ , the producer knows the distribution $h(\mu)$, and the producer and the consumers act so as to maximize their expected gain.

Integrating g , we obtain the cumulative distribution $G(\mu; w) = 1 - e^{-\mu w}$. From Eq. 6 we can compute the conditional distribution

$$\begin{aligned} \tilde{g}(\mu; w) &= \frac{g(\mu; w)}{1 - G(\Delta)} \Theta(w - \Delta) \\ &= \mu e^{-\mu(w - \Delta)} \Theta(w - \Delta). \end{aligned} \quad (11)$$

The average valuation for this conditional distribution is $\bar{w}(\mu, \Delta) = \Delta + \mu^{-1}$. Using Eqs. 5 and 9, we obtain the expected number of purchased articles and the expected surplus (assuming the consumer subscribes):

$$\begin{aligned} \langle k(\mu) \rangle &= N e^{-\mu \Delta} \\ \langle s(\mu) \rangle &= \frac{N}{\mu} e^{-\mu \Delta}. \end{aligned} \quad (12)$$

To compute the producer's expected profit as a function of Δ and F , we can substitute Eqs. 12 into Eq. 10, which yields:

$$\begin{aligned} \langle \Pi(F, \Delta) \rangle &= \frac{1}{(\mu_{\max} - \mu_{\min})} \int_{\mu_{\min}}^{\mu'} d\mu (F + N(\Delta - \gamma) e^{-\mu \Delta}) \\ &= \frac{(\mu' - \mu_{\min})F + N(1 - \frac{\gamma}{\Delta}) (e^{-\Delta \mu_{\min}} - e^{-\Delta \mu'})}{(\mu_{\max} - \mu_{\min})}, \end{aligned} \quad (13)$$

where μ' is defined as $\mu' = \min(\mu_{\max}, \tilde{\mu})$ and $\tilde{\mu}$ is defined as the unique solution to $N e^{-\tilde{\mu} \Delta} = F \tilde{\mu}$.

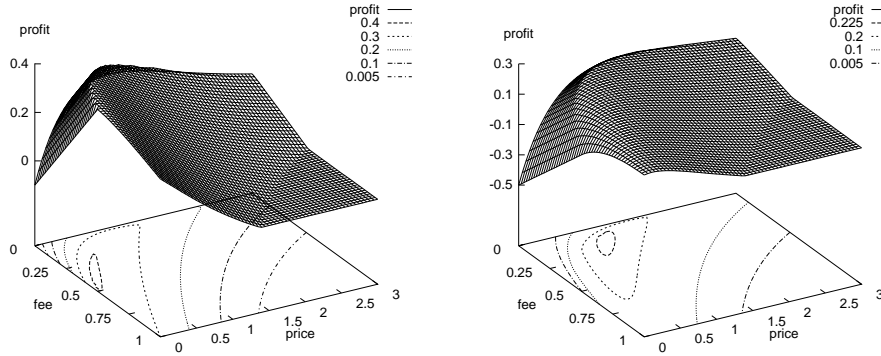


Fig. 1. Profit landscape $\pi(f, \Delta)$. h is a uniform distribution with $\mu_{\min} = 0.5$, $\mu_{\max} = 2.0$.
a) Production cost $\gamma = 0.1$. b) Production cost $\gamma = 0.5$.

Eq. 13 can be visualized as a profit landscape in which the expected profit is plotted as a function of F and Δ . It is convenient to define a normalized expected profit $\pi = \Pi/N$ and a normalized fee $f = F/N$. Fig. 1 illustrates the landscape for two different production costs: $\gamma = 0.1$ and $\gamma = 0.5$.

In each such landscape, there are two ridges. The lower ridge, which demarcates the boundary beyond which the producer attracts no consumers and thus makes no profit, is defined by $e^{-\mu_{\min}\Delta} = f\mu_{\min}$. The upper ridge is described by a piecewise joining of two nonlinear curves, the simpler one being described by the relation $e^{-\mu_{\max}\Delta} = f\mu_{\max}$. This portion of the ridge has the following characteristics:

- it is formed by a discontinuous derivative with respect to f and Δ (it changes abruptly from positive to negative)
- for price settings along this ridge, all consumers subscribe
- if the production cost $\gamma < \gamma_{\text{crit}}$, the optimal setting of (f, Δ) occurs along this part of the ridge. For the examples of Fig. 1, the critical production cost γ_{crit} is found to be approximately 0.28⁴, so that Fig. 1a ($\gamma = 0.1$) is in this regime, while Fig. 1b ($\gamma = 0.5$) is not.

The other portion of the ridge is defined by a more complex nonlinear relation between f and Δ . It is less sharp, resulting from derivatives with respect to f and Δ being zero rather than jumping discontinuously from positive to negative.⁵

⁴ It can be shown that, for exponential g and uniform h , $\gamma_{\text{crit}} = \frac{\Delta(\gamma_{\text{crit}})\mu_{\min}\mu_{\max} + \mu_{\min}}{\mu_{\max}^2} \geq \frac{\mu_{\min}}{\mu_{\max}^2}$, which is positive when $\mu_{\min} > 0$.

⁵ These characteristics only pertain to the case where g is exponential and h is uniform; we have explored other combinations of functional forms for g and h that yield profit landscapes that are topographically different. For example, if g is exponential and

Figure 2 shows the dependence of the optimal f and Δ upon the production cost γ . There is a discontinuous derivative at $\gamma = \gamma_{\text{crit}}$, due to the switchover between the two nonlinear curves that define the upper ridge in the landscape. As one might guess, the profit π decreases monotonically with the production cost. The proportion p of consumers that subscribe is 1 for all $\gamma \leq \gamma_{\text{crit}}$; for γ exceeding this threshold the proportion of subscribers is strictly less than one, and is given by $\frac{\hat{\mu} - \mu_{\min}}{\mu_{\max} - \mu_{\min}}$.

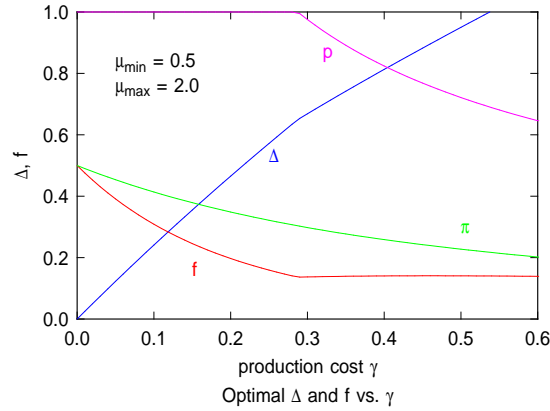


Fig. 2. Optimal Δ , f , π , and p vs. γ , where h is a uniform distribution with $\mu_{\min} = 0.5$, $\mu_{\max} = 2.0$.

4.2 Uninformed consumers

For a variety of reasons, a consumer may have to rely on adaptive estimates of its valuation distribution, g . Suppose for example that consumer i knows that g is an exponential distribution, but does not know its individual parameter μ_i . Depending on the consumer's beliefs about the dynamics of the environment, it might wish to place more or less weight on recent observations. Since the mean of the distribution $g(\mu; w) = \mu \exp^{-\mu w}$ is $1/\mu$, one reasonable and flexible approach to estimating μ is to start with an assumed prior μ_0 and after each period of subscription to update the estimate $\hat{\mu}$ according to

$$\hat{\mu}_{t+1} = (\phi * m_t + (1 - \phi)(\hat{\mu}_t)^{-1})^{-1}, \quad (14)$$

where m_t is the mean of the N valuations received during the subscription period t , and ϕ is the consumer's "flightiness" factor. This factor could be set to

h consists of a number of well-separated mass points, the landscape can contain multiple ridges and peaks.

a constant, or be time-varying; in fact, by setting $\phi = 1/t$, the consumer weights all observations equally.

Suppose that the producer knows the distribution h , and sets f and Δ to the “optimal” values given by Fig. 2. (For $\gamma = 0.1$, these values are $(f, \Delta) = (0.30878, 0.24098)$, at which the “ideal” profit $\pi = 0.4137$ and the fraction of subscribers $p = 1$.) Furthermore, suppose that the consumers all start with exactly the right estimates of their μ parameters, but that they update their estimates in accordance with Eq. 14. Fig. 3 illustrates what happens to consumer profits and subscription rates as a function of time for two values of ϕ . In both cases, the proportion p of subscribed consumers continues to diminish indefinitely, and, correspondingly, so does the profit π .

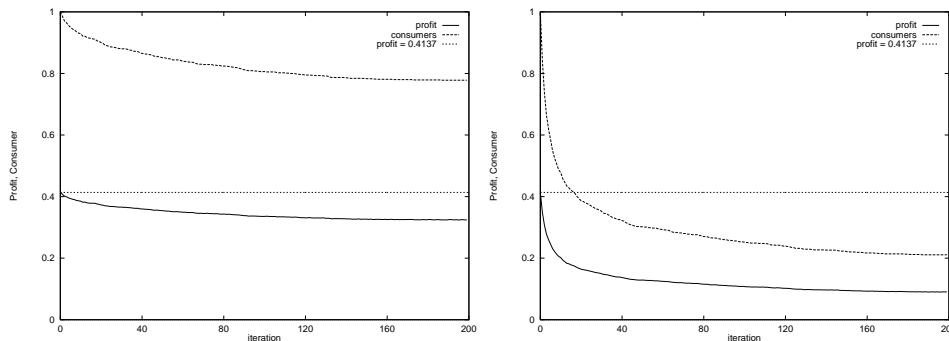


Fig. 3. Profit π and proportion of subscribed consumers p vs. time (in subscription periods) when consumers estimate their μ values, and the flightiness parameter is a) $\phi = 0.1$ and b) $\phi = 0.9$. Consumer leakage occurs for $\phi = 0.1$, and is even more pronounced for $\phi = 0.9$. Profits diminish indefinitely, although the rate of reduction slows with time. The market consists of $M = 1000$ consumers and one seller offering $N = 10$ articles per subscription period. Dashed horizontal line indicates ideal profit of 0.4137.

Fig. 4 offers further insight into consumer leakage. It shows the expected average profit over 200 subscription periods for a wide range of f and Δ (not just the optimal value). These are represented as profit landscapes for $\phi = 0.1$ and $\phi = 0.9$. Both landscapes have become peaked at lower values of f . In other words, if the producer insists on holding to a fixed price schedule, it will optimize profits by setting the subscription fee lower than it could charge perfectly informed consumers, thereby encouraging overly pessimistic consumers to stay in the market longer. It partially compensates by raising Δ somewhat.

A little thought suggests that, regardless of γ , ϕ , or other such parameters, consumer leakage will eventually lead to complete market failure for any fixed setting of f and Δ . As consumers continually update their $\hat{\mu}$, they are in effect conducting a random walk about the correct value of μ . If it should ever happen that the consumer’s $\hat{\mu}$ rises above a value such that $\langle s(\mu, \Delta) \rangle$ falls below F , the

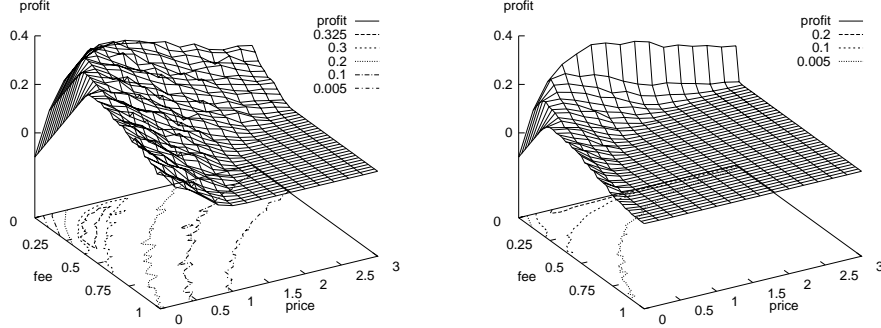


Fig. 4. Profit landscape after consumer leakage has occurred. Mean profit per article per consumer as a function of F and Δ , after 200 subscription periods with F and Δ at the values that would be optimal for perfectly informed consumers. The market consists of $M = 1000$ consumers and one seller offering $N = 10$ articles per subscription period. a) $\phi = 0.1$. b) $\phi = 0.9$.

consumer will become permanently disenfranchised from the market. In other words, the value of $\hat{\mu}_{\text{crit}}$ at which consumers are discouraged from subscribing (which, from Eq. 12, can be computed from $\frac{1}{\hat{\mu}_{\text{crit}}} e^{-\hat{\mu}_{\text{crit}} \Delta} < f$), acts as an absorbing boundary. In the next subsection, we will discuss how a producer might use dynamic pricing and optimization to prevent unchecked profit erosion.

4.3 Solving the leakage problem

Consumer leakage hurts both the producer and the consumers, and therefore all players have an incentive to counteract it. Both consumers and producers can do this by putting more emphasis on exploration as opposed to (pure) exploitation. Consumers could use a variety of schemes; for example, they could choose to subscribe at random with a finite probability even if their expected surplus is less than the subscription fee, and this probability could diminish monotonically as the difference between these quantities increases. Producers could fight leakage by temporarily decreasing prices to resurrect consumers who have mistakenly disenfranchised themselves, in hopes that with additional samples the consumers will increase their estimated surplus to levels that can support higher prices. One might expect that consumers could be enticed back into the market even with small discounts, since, once their $\hat{\mu}$ ventures into a realm where subscription appears to be unprofitable, it is frozen at this just-barely-unprofitable value.⁶

⁶ Economists have studied behavior of this sort in order to explain temporal price dispersion, such as “sales” and temporary discounts. A mixture of reasons have been modeled, including attempts to price discriminate between better and worse

Here we focus on the producer’s strategy for enticing overly pessimistic consumers back into the market. (This is not to deny that consumers’ exploration strategies merit serious study.) From our study of consumer leakage, it is apparent that the producer’s strategy must involve dynamic pricing, and that it must cope with a profit landscape that changes dynamically due to shifts caused by consumers’ ongoing attempts to learn an estimate of g . It also seems most likely that the pricing strategy would involve stochastic search on this dynamically changing landscape, rather than following some pre-planned schedule.

The dynamically changing nature of the profit landscape in our problem limits the applicability of standard stochastic optimization techniques. For example, approaches such as simulated annealing implicitly assume that the search is being conducted on a static landscape as the value of its temperature parameter is lowered. While more sophisticated approaches that are currently oriented towards static landscapes might be modified to dynamic ones, we focus here on a very simple iterative random-mutation hill-climbing method (RMHC). At each iteration of the RMHC method (i.e. following each subscription period), the producer tries a new f and a new Δ . If the new settings result in increased profit, the producer continues to move in the same direction in the f - Δ landscape in the next trial. Otherwise, it backtracks to its previous best point and makes small adjustments to f and Δ . The step size of the adjustment is chosen from an exponential distribution with a mean $1/\alpha$ for both f and Δ . To take into account the dynamic nature of the landscape being maximized, the hill climber has finite memory and can remember the previous best point visited within the last T time steps. We have observed that, for overly large values of T , the producer can get mired for a long time in regions of the (f, Δ) landscape that have become less profitable due to evolving consumer expectations.

Fig. 5 illustrates how the profit π and the subscribed consumer population p evolve with time for a particular setting of $\alpha = 0.5$ and $T = 20$. For these settings (and for a fairly broad range around them), the RMHC algorithm prevents long-term consumer leakage and profit erosion, even when the consumers are extremely flighty ($\phi = 0.9$). Frequent lowering of prices continually pulls consumers back into the market so they can keep sampling articles. Ironically, the RMHC’s noisy exploration of the profit landscape helps to keep that landscape relatively stable. While the profit never attains the value of 0.41 that it would have if the consumers were to keep their $\hat{\mu}$ ’s fixed at their exact values, the producer is still able to extract a healthy, reasonably stable profit of roughly 0.25. Fig. 6 illustrates the noisy trajectory of f and Δ .

5 Conclusions

Commerce in information goods is one of the earliest emerging applications for intelligent agents in commerce. However, the fundamental characteristics of information goods mean that they can and likely will be offered in widely varying

informed customers, and inducing potential customers to bear some search costs to find a better price or product. See [11], [10], and [12].

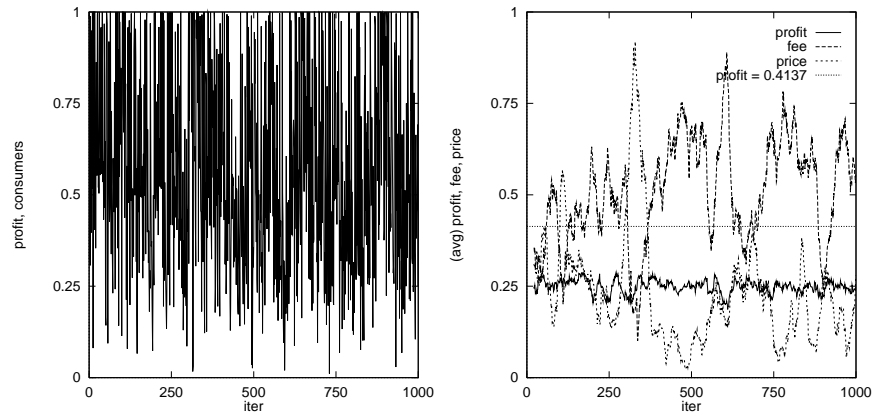


Fig. 5. Simulation run with producer using RMHC algorithm with memory of $T = 20$ subscription periods and step-size parameter $\alpha = 0.5$. $\gamma = 0.1$, $\phi = 0.9$. a) Proportion p of subscribed consumers. b) Moving average of f , Δ and π (integrated over 20 subscription periods).

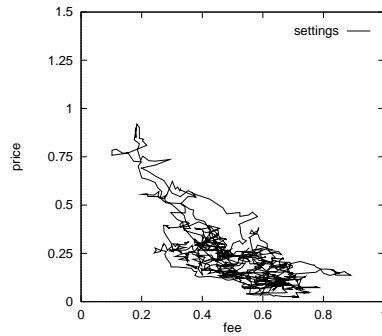


Fig. 6. Trajectory of f and Δ taken by the RMHC algorithm for the simulation run shown above.

configurations. Participating agents will need to deal with uncertainty about both prices and location in multi-dimensional product space. Thus, studying the behavior of learning agents is central to understanding and designing for agent-based information economies. Since uncertainty will exist on both sides of transactions, and interactions between learning agents that are negotiating and transacting with other learning agents may lead to unexpected dynamics, it is important to study two-sided learning.

We presented a simple but powerful model of an information bundling economy with a single producer and multiple consumer agents. We then explored the pricing and purchasing behavior of these agents when articles can be bundled. In this initial exploration, we studied the dynamics of this economy when consumer agents are uninformed about the distribution of article values. We discovered that a reasonable albeit naive consumer learning strategy can have a profound influence on market behavior — in this case, a strikingly bad influence.

Our consumer and producer agents were rather naive in our first learning experiments. This could be viewed as a criticism of our modeling. However, especially early in the development of adaptive agent intelligence, it may well be that agent-based markets are quite vulnerable to odd behavior and dysfunctional dynamics of the sort we observed. Our consumer agents did not recognize the option value of new information, and thus suffered by not undertaking sufficient exploration relative to exploitation. Our producer agent did not initially adapt to the pathological dynamics induced by the consumer agent naivete, and thus suffered by relying too confidently on its “perfect” but static knowledge. Although it was fairly easy for us to see what was going wrong, and to modify the producer agent in a simple way that ameliorated much of the problem, our environment is artificially simple and static. In more realistic settings it may be quite difficult for even relatively intelligent agents to adapt to emergent pathologies. Human markets may not be as susceptible because human behavior is less rote and more reflective. The lesson for agent design is to search for strategies that are dynamically robust and adaptive in the face of substantial uncertainty.

We have started to explore how to make our simple mechanism more robust in realistic settings. For example, the hill climbing algorithm is likely to get stuck at local optima. For the uniform h there is a single peak in the profit landscape, but that is not at all general. Several powerful optimization techniques, such as simulated annealing, exist for static landscapes. Extensions to these that handle dynamically changing, noisy landscapes, may lead to robustly adaptive agent learning strategies.

We have an active agenda of continuing work on this topic. For example, we have begun to consider less naive consumer strategies, that balance exploration against exploitation. We are considering producer strategies that adapt based on the number of recent subscribers relative to the producer’s model of the optimal number of subscribers. Perhaps most challenging but essential to a more general understanding of the problem is the extension of our work into an economy with multiple producers who are underinformed about each other’s competitive strategies as well as about consumer valuations.

In this and earlier work we have found that initial plausible but simple designs of economically-intelligent agents lead to dynamic market interactions that can be surprising and unsuccessful. The value of intelligent agents in electronic commerce will depend on the ability to understand the problems of learning and adaptivity, and to design agents that interact robustly in the presence of substantial uncertainty about both parameters and the strategies of other underinformed agents.

References

1. Y. Bakos and E. Brynjolfsson. Bundling information goods: Pricing, profits and efficiency. In D. Hurley, B. Kahin, and H. Varian, editors, *The Economics of Digital Information Goods*. MIT Press, Cambridge, Massachusetts, 1998.
2. J. C. Chuang and M. A. Sirbu. Network delivery of information goods: Optimal pricing of articles and subscriptions. In D. Hurley, B. Kahin, and H. Varian, editors, *The Economics of Digital Information Goods*. MIT Press, Cambridge, Massachusetts, 1998.
3. S. Fay and J. K. MacKie-Mason. Competition between firms that bundle. Manuscript, Dept. of Economics, University of Michigan, 1998.
4. J. E. Hanson and J. O. Kephart. Spontaneous specialization in a free-market economy of agents. In *Proceedings of Workshop on Artificial Societies and Computational Markets at the Second International Conference on Autonomous Agents*, May 1998.
5. J. O. Kephart, J. E. Hanson, D. W. Levine, B. N. Grosz, J. Sairamesh, R. B. Segal, and S. R. White. Dynamics of an information-filtering economy. In *Proceedings of Second International Workshop on Cooperative Information Agents (CIA-98)*, Paris, July 1998.
6. J. O. Kephart, J. E. Hanson, and J. Sairamesh. Price-war dynamics in a free-market economy of software agents. In *Proceedings of Alife VI*, June 1998.
7. J. K. MacKie-Mason and H. R. Varian. Some economics of the internet. In W. Sichel, editor, *Networks, Infrastructure and the New Task for Regulation*. University of Michigan Press, Ann Arbor, Michigan, 1996.
8. A. Mas-Colell, M. Whinston, and J. Green. *Microeconomic Theory*. Oxford University Press, 1995.
9. W. Oi. A Disneyland dilemma: Two-part tariffs for a Mickey Mouse monopoly. *The Quarterly Journal of Economics*, 85(1):77–96, February 1971.
10. S. Salop and J. Stiglitz. Bargains and ripoffs: A model of monopolistically competitive price dispersion. *Review of Economic Studies*, 44:493–510, October 1977.
11. Y. Shilony. Mixed pricing in oligopoly. *Journal of Economic Theory*, 14:373–388, April 1977.
12. H. R. Varian. A model of sales. *American Economic Review*, 70:651–659, September 1980.
13. J. M. Vidal and E. H. Durfee. Learning nested agent models in an information economy. *Journal of Experimental and Theoretical Artificial Intelligence (special issue on learning in distributed artificial intelligence systems)*, 10(3):291–308, 1998.
14. J. M. Vidal and E. H. Durfee. The moving target problem in multi-agent learning. In *Proceedings of the Third International Conference on Multi-agent Systems (ICMAS98)*, 1998.