Opening Up the ACM Digital Library:
An Alternate Method of Payment for the ACM Portal

Kamal Jain

Motivation: First, I would like to thank the Complexity Blog for this opportunity to publish my thoughts. The first time I wrote something here my objective was fulfilled. I asked for a free electronic version of "the algorithmic game theory book" at http://weblog.fortnow.com/2006/01/free-electronic-editions-of-new.html and we got it at http://www.cambridge.org/journals/nisan/downloads/Nisan_Non-printable.pdf. I have met people who were not sufficiently interested in this subject to buy this book for money. But they accessed the book in the free electronic format, got interested in the subject, and then purchased a hard copy. So a happy ending with big thanks to Lance Fortnow and Bill Gasarch, together with everybody who commented on the post including Lauren Cowles of Cambridge University Press.

As the title of this post suggests, I am making a similar proposal for the ACM portal. Some of us are fortunate enough to be in big universities and institutions, which pay for the ACM portal. But not all of us are equally fortunate. Sometimes even I am not either. When I access ACM Portal from home I am not fortunate enough unless I am on my office computer and willing to create a VPN connection to my office, which is often enough pain for me to postpone the activity to the next working day. I wonder if there are people who could learn something from the invaluable material at ACM portal, if it were available for free. I bet yes. And I bet they are in big numbers. Just like the number of people consulting an encyclopedia has grown up since Wikipedia is available. **The main problem I want to solve is how to make ACM Portal freely available to those people, who neither belong to subscribing institutions nor are subscriber themselves, but could gain invaluable knowledge if ACM portal were available for free.** There are associated costs of editing journals and also running the portal. So any proposed solution must recover these costs. **Beyond that, IMHO, it is our moral duty as true scientists to maximize the spread of the knowledge we create.**

Objective: People who know researchers at Microsoft Research know that academic scholarship is central here, in addition to efforts in support of product R&D. For the last few years, I have been thinking about how to make the scientific content available for free. We all know that the scientific content is not written for making profits but written for the benefit of the society at large. So is not it natural to find ways to make ACM portal available for free?

Knowing that I think about these issues, Eric Horvitz, a Microsoft researcher and the president of AAAI society (http://aaai.org, Association for the Advancement of Artificial Intelligence, one of the largest scientific membership organizations of AI researchers) engaged me about related issues. Here's a portion of a note he sent me:

> *These ideas came in the context of a discussion of AAAI's exploration of new access models for its scholarly content at a discussion at the AAAI Executive Council meeting*

last summer. How might the costs of providing enhanced access be covered? We focused on the challenge of addressing our desire as a society to enhance access to AAAI content, but to also support the costs of doing so. To date, our content is available to AAAI members who pay an annual membership fee—and this is one of the benefits of being a member. Access to content and the indexes are also provided to institutions such as libraries for fees. Broadening the access of content resonates deeply with our goals as a scientific society to stimulate scholarship and collaboration. However, some are concerned that opening up access will lead to a loss of revenues from members and institutions, as the incentive to pay for the access will go away. At our meeting we reviewed the costs and benefits of different models for providing wider access. An interesting line of thought was introduced by Tuomas Sandholm at the meeting: AAAI (and other scientific societies) are expert at providing high-quality peer review and organizing energetic scientific communities of people producing content and receiving reviews, feedback, and certification by the society. The people providing web access, e.g., via search engines (or other portals) are expert at providing the access. Might AAAI and other scientific societies make some trade with organizations providing access—might mechanisms be established more broadly for supporting reasonable trades?

I personally share the viewpoint of this note and like to have an open discussion on various mechanisms for providing access to scientific content, esp. those which can help make the scientific content available for free or at a deep subsidy. For the purpose of this blog discussion, let us describe the eventual goal as **free access to scientific content**. I had proposed a number of solutions for this problem in the past based on my **atomic economics theory, which states that the internet and the computer allow us to make decisions at their finest granularity thereby enabling us to make more optimal decisions, e.g., buying a single song instead of an entire album.** So when I saw that other researchers are also thinking about it, I naturally got excited. I was also fortunate this summer to have a fantastic economics intern, Alex White, from Toulouse School of Economics. We analyzed a number of solutions with the **goal of maximizing the access to the scientific journals**. We believe open non-exclusionary is as important as free, so we **paid more attention to those solutions which are non-exclusionary.**

Advertisements: One solution some people immediately put forward is the advertisements, i.e., the users pay for the portal by paying small accidental attention to the ads instead of paying dollars. Fortunately ACM portal already runs the advertisements. **Unfortunately advertisements are not sufficient or suitable for the content which is highly valuable to only a small community of user**. Let us look at Yahoo's annual finances. Wikipedia mentions that it has $7.5 Billion in annual revenue and an audience of 1.5 Billion, therefore only $5 value per person annually. **That's not sufficient amount of money to pay for our scientific journals.** The amount of ad money for a scientific portal would be even lower because its audience would be focused and less distracted to ads. **So the content which is highly valuable to a small community currently has to be funded by the large payments from its community of the users, thereby restricting its use to the most devoted core of the community**.

Commercial Deals: ACM runs a discount programs for its members at http://www.acm.org/membership/discounts. It is not uncommon to receive a credit card solicitation or insurance solicitation from ACM or on behalf of ACM. Given that these discounts are commonly available by mentioning any number of promotional codes available on the web, at least ACM must be receiving some money for providing our contact information to these credit card or insurance companies. Given that these companies belongs to old-economy or pre-internet era, they do not have enough profitability or efficiency to completely pay for our ACM subscription. Take the example of WorldPoints credit card advertised on the ACM homepage. This credit card already pays 1 Point for every dollar spent to its users irrespective of whether the user is an ACM member or not. This is due to the competition in the credit card business. So if the credit card could have paid our annual ACM portal subscription fee then it would be better off paying that amount to its users directly, irrespective of whether a user is an ACM member or not. Of course, there is a tiny bit of influence ACM has on its member's commercial decisions, so ACM gets a tiny bit of fee from its advertising partners.

One of the proposals we considered is to take this existing program and apply it to the internet age in those domains where both the profits and the ACM influence are high enough to pay for our entire ACM subscription. Before proposing such a domain let me ask the readers a trivia question: Suppose a new $500 computer is bought today for home or education environment. Which company's R&D is likely to be paid the most over its lifetime of say 5 years? (I am restricting to R&D to avoid physical consumables such as electricity, optical fiber, silicon wafer etc., which is charged to the Mother Nature, who can't pay for our ACM subscription.)

The answer might surprise some of the readers of this blog. **The answer is likely to be a search engine company**. Search engines is as big a revenue source as, say, operating systems (if some people guessed that), but the majority of the operating systems revenue is generated on the enterprise computers whereas the majority of the search engines revenue is generated on the home computers. Therefore on a home computer, search engines are likely to outdo operating systems and also the R&D part of other products.

So the search engines industry is potentially a right candidate to make commercial deals. The search engine industry satisfies some other nice properties too:

1. Even though the search engine industry is paid the most, **users do not pay for it directly**. That means the search engine industry is non-discriminatory or perfectly discriminatory depending upon the viewpoint. Non-discriminatory because everybody, rich or poor, has equal access to a search engine. Perfectly discriminatory in the (economics) sense that the search engine industry has perfect price discrimination. A user pays as much as he/she could pay by paying accidental attention to the ads.

2. ACM could have a high influence on its community members in their choice of default search engines. Certainly a much higher influence than the choice of the credit card.

3. Even though there are multiple search engines, most search engines are highly profitable on marginal cost basis (of another search). So the more the volume they have, the more profitable they could become.

4. Search engines and the open content (i.e., web) are complementary products. Everybody likes a library with great books and a great index. Suppose you have a library with a great index but no books. Would this library be useful? Suppose you have a library with great books but no index. Would this library be fully useful? Just like great books and a great index complement each other, search engines and the open content complement each other. In economics, whenever you have a complementary set of products, **the best use of these products occurs when there is a tie-in between these complementary products**. Tie-in of two complementary products is typically better for everybody, i.e., not only better for the users but also better for the providers of both the products, w.r.t. their objective functions.

   Let us look at an example of a pair of complementary products. Cell phone and cell service complement each other. One is useless without the other. At least in the United States, which has a robust framework of enforcing contracts, cell services subsidize cell phones, and often to the extent of free. One explanation for this subsidy is that people find it easier to pay for cell services, which is naturally is in monthly installments, than to pay for cell phones which is a large upfront cost. The same could happen between search engines and open content. People find it easier to pay for the search engines by letting search engines show them ads than to pay for the content.

   Assuming that we agree on search engines industry being a good candidate to seek commercial contracts to help pay for the scientific content, the next question is what kind of contracts best serve the goals of scientific community which search engines would also be willing to sign.

Types of Contracts: Let us look at some types of candidate contracts. Please feel free to suggest more candidate contracts in the comment section.

1. **Exclusive Contract**: This contract means that a scientific society sells the annual rights exclusively to a single search engine for a suitable price with the condition that the content is provided to their users for free. This is similar to selling broadcast rights of a sporting event to a TV channel, which is free to the audience. The main advantage of this type of contract is that the scientific society could get a premium price in exchange of granting exclusivity. If we take our phone example, iPhone perhaps gets a bit higher subsidy than other phones because of granting exclusivity to AT&T.

   Besides my ideological concerns against granting exclusivity in a social context, there are some disadvantages too in granting exclusivity. The search engine market is way less competitive with a clear dominating search engine, Google. Google buys some content for Google News from AP

(http://ap.org, Associated Press) on a non-exclusive basis, but Google is not a dominating firm in the news category. In the search business where it is dominating, it may not want to start a culture of paying for the content.

On the flip side, even if Google is willing to pay for the scientific content, we scientists should not sell our highly valuable content exclusively to it. Doing so will further consolidate search engine industry in Google's hand so that it may decline to pay for the scientific content after few years. In other words this solution is not perpetual.

How's about selling the content exclusively to other search engines? Due to exclusivity other search engines will find higher value in the content, which is a good thing. But then, **if I point to my academic hat I am wearing now**, I still raise ideological concerns. Also, in the long term the market may split. One search engine may buy ACM content and another may buy IEEE content. Therefore a scientist would need to access multiple search engines to find the content and not built habit of any. In other words, no search engine may find the value of exclusivity to pay a premium for it. So a scientific society may end up granting exclusivity without any additional premium.

2. **Pay by the usage**: In this solution no search engine gets exclusivity. Any search engine could show the links to scientific content. Search engines which are paying a fee, or in other words sharing the advertisement revenue with the scientific society get premium experience for their searchers. If the search engine is not paying any fee, then clicking a link to an article shows only the bibliography information with abstract, and a login screen. But if the search engine is paying a fee then the same link shows the full article. One can find this experience when accessing the premium content of WSJ.com through Google News. Clicking the WSJ.com link from Google News show the full article whereas clicking the same link from WSJ.com homepage or other news portals shows only the abstract and a login screen.

This is an ideal solution for the long term. Just like advertisers pay a small fee for a click, here the scientific society gets paid a small fee for a click. So in essence, each search engine pays proportional to the traffic it derives by showing the full scientific content for free. In essence the trade is open and happening at its finest granularity (a principle of atomic economics).

Since the search engine market is not fully competitive, it is hard to achieve this ideal solution in the short term. The leading search engine may not want to start the culture of paying for content. A trailing search engine would not find as much value, since the users may search on the search engine only for the scientific content and not pay any attention on the ads. Even if by some changes to user interface, the search engine place ads at positions where the users see the ads, it is still not very useful because the intent of the user is much focused and that is to discover the free scientific article. So very few advertisers would pay a high enough rate for such ads for the search engine to afford pay for the scientific content. But in the long term, this

would work because the search engines which provide only the abstract and not the full articles would lose their searchers to other search engines.

3. **Cobranded search default or toolbar**: In the short term, just like ACM has a cobranded MasterCard with Bank of America (https://wwwa.applyonlinenow.com/USCCapp/Ctl/entry?sc=LYH1), ACM could have a cobranded search default/toolbar with search engines. Again wearing my academic hat, I like to recommend it on the non-exclusive basis. First let me define "search default". A search default is basically setting in a web browser a default search engine. Most modern browsers have a notion of search default. Similarly most toolbars have a search field built in. A co-branded search default could be ACM+Google or ACM+Yahoo. By non exclusive, we mean that ACM provides an API so that any search engine who is willing to pay for the content as in 2., could create a cobranded search default. ACM permits multiple search engines to build their cobranded search defaults. The main difference here from 2 is that a search engine gets some stickiness. A user besides running the searches for the free scientific content may also run other regular searches, with well paying ads. By having a toolbar on the client machine, a participating search engine may also have a non-zero probability of correctly estimating the stickiness factor. I like to refer the readers to contract theory to see that the payment to the scientific society is tied to those factors which could somehow be estimated.

Assuming that our ACM community could successfully negotiate a deal with a search engine, the next question is how much a search engine needs to pay? Do they have to pay $10 per article as the current retail price or just 10 cents or even 1 cent?

Economics of the Content: Let us assume that serving the content has almost zero marginal cost. Suppose a scientific society wants to figure out the price it needs to charge for its content. Assume that the scientific society has a fixed cost of C. For conceptual simplicity, assume that the scientific society only retails the document and sells no subscription. The scientific society looks at the demand function $D(.)$, and finds the minimum p so that $p \times D(p) = C$. For simplicity of the example, assume that such a p exists. Clearly p would be a positive number. Now suppose the scientific society is selling the content to a search engine who promises to provide it for free, then how much the society has to charge the search engine for each click? Under conceptual simplicity, the answer if $C/D(0)$. If $D(0)$ is very large compare to $D(p)$ then the search engine would have to pay a small price per use. There are several reasons for $D(0)$ to be large. For an example, if it is convenient and zero cost to rediscover an article then the same article would be repurchased multiple times. Of course, this is only a conceptual explanation. In reality, the pricing details could be complicated, e.g., the search engine may cap the payments per user to avoid fraud.

This is getting long, so only the deeply interested are likely to have gotten this far. If you have, then there might be several questions in your mind. Please add them to the blog discussion. Here, I am answering some of the questions which academic community including both researchers and professors has already risen.

FAQ:

Q1: Search engines pay to computer manufacturers for setting search default, and if the search engines pay a similar amount then by reasonable estimates, that amount may not be enough to pay for the ACM portal.

A1: I agree that that amount may not be sufficient. But the payments to ACM portal could be higher. When you buy a new computer the computer manufacturer does not have the last word on what search engines you would use. You have the last words. So if a computer manufacturer makes a choice for you then there is a certain probability that you would ignore the choice. So the payment to the computer manufacturer has to account for this probability. When you make the choice, by definition, you are using it. When you change the choice, you may stop getting the benefit too. So on per time basis, if you as a user negotiate your choice with the search engine, you may get better benefits such as the free access to ACM portal.

Q2: Would not it mean that I have to compromise the use of my favorite search engine?

A2: Not really. If this proposal is effective then your favorite search engine would be compelled to offer you the same benefit. Just look at your credit card benefit program, you may see free insurance, extended warranty, and a reward program. Pretty much any credit card company is compelled to offer these benefits, since they are desired by the users.

Q3: What's about the short term until this proposal demonstrates its effectiveness in attracting users?

A3: In the short term you may want to make one of the participatory search engines as your default choice and use your favorite search engine as a fall back choice. The tradeoff would depend upon how much value you give to free ACM content vs. the difference in value you consider between your favorite search engine and a participatory search which gives you free ACM content.

In the short term you could also make a personal choice of working with the status quo, i.e., not get the free ACM content and keep your favorite search engine as your default choice.

Q4: How long would it take from the short term to the long term situation where my favorite search engine likes to participate too?

A4: Ideally we should hope that your favorite search engine decides to participate in the first round itself. Today major search engines are run by companies who have a good relationship with scientific communities. So they may want to participate at the very beginning to strengthen their relationship by supporting the objectives of the scientific communities. In case your favorite search engine decides to not participate until compelled then the time to reach a long term equilibrium will depend upon how soon other organizations, such as sister scientific organizations, with closed content open up their content for such an exchange of value. The time to equilibrium will also depend upon how much support the communities, represented by those organizations, extend to this value exchange. Good news is that the internet time runs fast, that should help shorten the time to long term equilibrium.

Q5: Does not payment between a search engine and content owners (websites) could interfere with the search neutrality?

A5: Search engines already have two sections. First, where they display openly available content and claim to be unbiased. Second, where they show ads and charge money for the clicks. There is no obvious social reason for them that they can't have a third section, where they show desirable content and pay for it, in order to attract more searchers. Since they claim to do an unbiased job in the first section in the presence of the advertising links, they could try to do the same unbiased job in the presence of the sponsored content links. The search engines may be maintaining the current status quo, because it may suite their purpose. The status quo may not necessarily be the best thing for the society. We might have stuck in the status quo, because there is not enough competition in the search engine industry.

Q6: The web is currently free and open. That also has helped the growth of it. Does not payment exchange make the web close?

A6: The entire web is not currently free and open. I can't fully access the ACM portal. The content which is viable as free today would remain viable as free tomorrow. On the other hand, some of the content which is not yet viable as free today may actually become viable as free under the payments from the search engines. Some of the content might not have been created today because of the lack of any business model, which could take advantage of synergies between the complementary nature of the search engines and the web. A new business model which could only bring additional money to the web could only help grow the web further. For an example, search engines show ads next to the Wikipedia results and if they could donate a fraction of that ad revenue to Wikimedia org, then it can grow only further and not shrink. Or maybe other similar websites, who do not mind taking commercial money, could do a better job because they won't have to depend on lowly paid contextual ads, but could target a fraction of highly paid search ads. In order to negotiate money from the search engines, they have to optimize their websites to make them attractive to the users instead of optimizing for the advertisers. We may be missing on such opportunities.

 Q7: If not all the content is available on a single search engine, then it may become necessary to use the multiple search engines. Would not this be a pain?

A7: That's a good question and non-exclusivity might help. The best search engines would try to get all the available content which attracts the users. The content which does not attract the users would anyway remain free.

In case some of this concern becomes true, that won't be a bad thing. This concern is true today too. Different search engines surface some of the different links. So if one search engine fails to answer a web-query then one can try another. So it is worthwhile to have a variety of search engines. In any case, just having a uniform common denominator is less desirable than having the heterogeneous extensions to that denominator.