The Authorship Dilemma: Alphabetical or Contribution? (Extended Abstract)

Margareta Ackerman
Department of Computer Science & Engineering
University of California, San Diego, USA
maackerman@eng.ucsd.edu

Simina Brânzei
Department of Computer Science
University of Aarhus, Denmark
simina@cs.au.dk

"In the long history of humankind (and animal kind, too) those who learned to collaborate and improvise most effectively have prevailed." **Charles Darwin (1809–1882)**

ABSTRACT

Academic communities have adopted different conventions for ordering authors on academic publications. Are these choices inconsequential, or can they significantly impact individual authors, or even communities at large? We consider a game theoretic model to study allocation of credit to authors; in particular, we find that alphabetical ordering can lead to higher research quality, while ordering by contribution results in a denser collaboration network and more publications.

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Sciences - Economics

General Terms

Theory, Economics, Human Factors

Keywords

academic collaboration; coalitions; social networks

1. INTRODUCTION

Research has been recognized as a crucial economic activity, towards which developed countries allocate a significant fraction of their resources. Allocation schemes should incentivize research communities to operate at their best, but decisions regarding the allocation of research resources are sometimes made in an ad-hoc manner, with little theoretical or empirical justification of their long-term effects. In this paper, we investigate one of the core problems in this domain – namely, the allocation of credit for scientific work.

The allocation of scientific credit influences funding decisions, as well as tenure, promotions, and awards. Given the critical role that credit allocation plays in academia, surprisingly little is known about the effects of name-ordering conventions. What influence do ordering schemes have on individual authors, and more globally, on the research communities where they are applied?

The prominent ordering conventions are to list authors alphabetically or in descending order of contribution to the paper. Alphabetical ordering is known to result in the *Matthew Effect*, whereby readers tend to assume that more established authors deserve more credit; in addition, it benefits authors with earlier surname initials [2, 3, 6]. Tenure decisions, fellowships, and to some extent even Nobel Prize winnings are correlated with surname initials [3]. For such reasons, the American Psychological Association mandates ordering authors by their contribution.

Nevertheless, major disciplines such as mathematics, theoretical computer science, and some branches of economics have a long tradition of relying on alphabetical ordering. The American Mathematical Society states that *"Determining which person contributed which ideas is often meaningless because the ideas grow from complex discussions among all partners... mathematicians traditionally list authors on joint papers in alphabetical order."*

We formulate a natural game theoretic model of collaboration that allows the investigation of credit allocation schemes. Our model offers a compelling explanation for the phenomenon that alphabetical ordering can lead to improved research quality in some communities. In particular, alphabetical ordering encourages collaborators to match each other's efforts, resulting in improved projects, while contribution ordering can result in the completion of more research projects and a denser social network. Both phenomena have been observed empirically (see Brown *et al* [1], Laband and Tollison [4]).

2. ACADEMIC GAME MODEL

Let $N = \{1, \ldots, n\}$ be a set of agents. Each agent i has a budget of weight w_i, consisting of a set of coins $C_i = \{c_{i,1}, \ldots, c_{i,n_i}\}$; coin $c_{i,j}$ has a weight $w_{i,j} > 0$, with $\sum_{j=1}^{n} w_{i,j} = w_i$. A project of weight w can be solved either by one agent who invests a coin, or by a pair of agents, each of which invests one coin (the coins add up to at least w). An agent can participate in multiple projects by allocating different coins to each, the same pair of agents can solve multiple projects together, and a coin can only be used once. We note that two-authored papers represent a substantial fraction of the literature in many fields (see, e.g., Newman [5]).

Solving a project of weight w gives a reward $\mathcal{F}(w)$, which is divided among the authors of the project. We study games where the reward function is homogeneous and convex: $\mathcal{F}(x) = \alpha \cdot x^d$, where $d > 1$ and α are constants. Unless otherwise specified, the agents can solve projects that are arbitrarily hard (or easy), and for each weight w, there are infinitely many projects of this weight.

Finally, in each academic community there is a general perception of the significance of being the first or second author on a paper. Without prior knowledge about the specific paper or its authors, the relative contribution of each author on a two-authored paper is given by a fixed *contribution vector* $[\phi, 1 - \phi]$, where $1 > \phi \geq 1 - \phi > 0$. Thus the community assumes the first author contributed ϕ% and the second $1 - \phi$%. We say the ordering is *alphabetical* when $\phi = 0.5$ and *contribution-based* otherwise.

Given an academic game, a *coalition structure* CS is a partition of all coins, such that every coin $c_{i,j}$ of agent i is either a singleton project, or is paired with a coin $c_{k,l}$ of another agent $k \neq i$. Let CS_i be the set of projects that agent i contributes to. The *utility* of i is: $u_i(CS) = \sum_{P_j \in CS_i} v_i(P_j)$ where $\{P_1, \ldots, P_m\}$ are the projects solved under CS, $w(P_j)$ is the weight of project P_j, and

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

$$v_i(P_j) = \begin{cases} w(P_j)^d & \text{if } i \text{ completes } P_j \text{ alone} \\ \phi \cdot w(P_j)^d & \text{if } i \text{ is first author on } P_j \\ (1-\phi)\cdot w(P_j)^d & \text{if } i \text{ is second author on } P_j \end{cases}$$

3. INDIVISIBLE BUDGETS

We first study indivisible budgets, where each agent owns a single coin and can be involved in a single project. Indivisible budgets already highlight an interesting difference between alphabetical and contribution ordering: there exist natural settings in which alphabetical ordering encourages agents to match each others' efforts, leading to the completion of larger projects.

First, we define a coalition structure CS to be *pairwise stable* if for all $i \in N$, $u_i(CS) \geq \alpha \cdot w_i^d$, and for all $i, j \in N$, with $w_i \geq w_j$, either $u_i(CS) \geq \phi \cdot \alpha \cdot (w_i + w_j)^d$ or $u_j(CS) \geq (1-\phi)\cdot\alpha\cdot(w_i+w_j)^d$. That is, no agent i can improve by allocating his coin to a singleton project, and no pair of agents can deviate to a joint project. Notably, stable coalition structures are guaranteed to exist under alphabetical ordering, but not under contribution (for details see full version of the paper).

3.1 Research Quality

Alphabetical ordering can result in higher research quality than is possible under some contribution-based scheme. Given an instance of a game, the most difficult project that can be solved results from the combined effects of the two strongest agents. We call a project of this difficulty a *hard project* and illustrate how credit allocation determines what type of collaborations take place when there are two types of agents, *heavy* and *light*. The weights are normalized so that a heavy agent has weight 1 and a light agent has $\lambda \in (0, 1]$.

THEOREM 1. *Consider an academic game with indivisible budgets and two types of agents, light and heavy. Then every pairwise stable coalition structure has:*

(1) *Only same-layer collaborations when:* $\frac{(1+\lambda)^d}{2^d+(1+\lambda)^d} < \phi < \min\left\{\frac{2^d-1}{2^d}, \frac{1}{(1+\lambda)^d}, \frac{2^d}{2^d+(1+\lambda)^d}\right\}$

(2) *Only cross-layer collaborations when:* $\max\left\{\frac{2^d-1}{2^d}, \frac{1}{(1+\lambda)^d}, \frac{2^d}{2^d+(1+\lambda)^d}\right\} < \phi < 1 - \left(\frac{\lambda}{1+\lambda}\right)^d$

(3) *No collaboration when:* $\frac{2^d-1}{2^d} < \phi < \frac{1}{(1+\lambda)^d}$ *or* $\phi > \max\left\{\frac{2^d-1}{2^d}, 1 - \left(\frac{\lambda}{1+\lambda}\right)^d\right\}$

Note that the more difficult projects are solved under alphabetical ordering.

3.2 Free Riding

Alphabetical ordering has been described as unfair as it gives the same credit to all authors even when they do not contribute equally. Formally, if two agents allocate weights x and y to a joint project, the rewards should be proportional to their efforts, i.e. $\left(\frac{x}{x+y}\right)(x+y)^d$ and $\left(\frac{y}{x+y}\right)(x+y)^d$. The *fair contribution vector* for this project is uniquely defined as: $\mathcal{C} = \left[\frac{x}{x+y}, \frac{y}{x+y}\right]$; all other vectors exhibit free riding. The *free riding index* of each agent i, $\mathcal{L}_i$, is the (normalized) difference between the perceived and actual contribution. We show the worst case for free riding occurs under contribution-based ordering.

THEOREM 2. *Consider an academic game in which the agents have indivisible budgets of arbitrary sizes. Then the highest free riding index of any agent under alphabetical ordering is at most* $0.5^{1/d} - 0.5$*, while it can be as high as* $0.5 - 0.5^d$ *under some contribution-based ordering schemes. Moreover, the highest amount of free-riding that occurs in any project solved under alphabetical ordering is smaller than under contribution-based ordering.*

4. DISCRETE BUDGETS

The results from indivisible budgets carry over to the general model, where the agents have multiple coins and can work on many projects at once. In addition, we uncover several other phenomena. Our solution concept is pairwise stability for games with overlapping coalition structures (see e.g. Zick *et al* [7]). The agents have sensitive reactions to a deviation, i.e. non-deviating agents that are hurt by a deviation drop all projects with the deviators, while unaffected agents are neutral and maintain the existing projects.

4.1 Rotations

There are meaningful scenarios where agents can overcome the limitations of a fixed contribution scheme by using *rotations*. That is, agents collaborating on multiple projects agree that one of them is the first author on half of their projects, while the other is first on the remaining projects (regardless of their actual contributions).

THEOREM 3. *There exist academic games with discrete budgets and multiple identical coins such that for every* ϕ*, the maximum number of hard projects is solved in a pairwise stable equilibrium and no free riding occurs.*

4.2 Implications for the Social Network

It has been observed empirically that contribution-based ordering results in denser collaboration networks. This can be seen in our model when the agents have heavy and light coins, where light coins represent very little effort ("cheap talk") but can nevertheless improve the quality of a paper and increase the number of collaborations.

THEOREM 4. *Consider an academic game with discrete budgets, where each agent has several heavy and light coins, of sizes* 1 *and* ε*, respectively, such that* $0 < \varepsilon \ll 1$*. Moreover, the conference tier is* 1 *and each agent has more heavy coins than light coins. Then whenever* $\phi > \max\left(\frac{2^d}{2^d+(1+\varepsilon)^d}, \frac{1}{(1+\varepsilon)^d}\right)$*, every pairwise stable equilibrium solves the maximum number of projects and the average number of collaborators per agent is the highest possible.*

5. ACKNOWLEDGMENTS

We would like to thank Peter Bro Miltersen and Yair Zick for useful discussion. Margareta Ackerman was supported by the National Sciences and Engineering Research Council of Canada (NSERC). Simina Brânzei was supported by the Sino-Danish Center for the Theory of Interactive Computation, funded by the Danish National Research Foundation and the National Science Foundation of China (under the grant 61061130540), and by the Center for research in the Foundations of Electronic Markets (CFEM), supported by the Danish Strategic Research Council.

6. REFERENCES

[1] C. L. Brown, K. C. Chan, and P. Lai. Marketing Journal Coauthorships: An Empirical Analysis of Coauthor Behavior. *J. Mark. Educ.*, 28(1):17–25, 2006.

[2] G. Efthyvoulou. Alphabet economics: The link between names and reputation. *Journal of Socio-Economics*, 37(3):1266–1285, 2008.

[3] L. Einav and L. Yariv. What's in a Surname? The Effects of Surname Initials on Academic Success. *J. Econ. Perspect.*, 20(1):175–188, 2006.

[4] D. Laband and R. Tollison. Alphabetized Coauthorship. *Appl. Econ.*, 38(14):1649–1653, 2006.

[5] M. E. Newman. Who is the best connected scientist? a study of scientific coauthorship networks. In *Complex networks*, pages 337–370. Springer, 2004.

[6] M. van Praag and B. van Praag. The Benefits of Being Economics Professor A (and not Z). *Jena Economic Research Papers*, (07-048/3), 2007.

[7] Y. Zick, G. Chalkiadakis, and E. Elkind. Overlapping Coalition Formation Games: Charting the Tractability Frontier. In *AAMAS*, pages 787–794, 2012.