

Góp ý sau khi đọc bài “Đánh giá nghiên cứu cơ bản: Công bố quốc tế” – tiêu chí quan trọng nhất”

Trước hết xin nói ngay là tôi đồng ý với dự thảo nêu trong bài báo trên ở Tia sáng ngày 4/5/2007. Nghiên cứu cơ bản nhằm tìm ra những tri thức khoa học mới và có ý nghĩa, làm nền tảng cho các nghiên cứu ứng dụng hay nghiên cứu cơ bản khác. Giá trị của nghiên cứu cơ bản là giá trị toàn cầu, và do vậy việc công bố quốc tế phải là tiêu chí quan trọng nhất. Khẳng định điều này là rất cần thiết, vì đa số các nghiên cứu của chúng ta vẫn chỉ công bố trong nước, và ai cũng biết dù đã lâu ta không nói ra là có một khoảng cách lớn giữa việc công bố trong nước và công bố quốc tế.

Điều tôi muốn nêu ở đây nhằm bổ sung thông tin sau: Vai trò của các bài báo ở tạp chí và hội nghị ở các ngành khoa học khác nhau là khác nhau, thậm chí rất khác nhau. Ở đây tôi giới hạn chỉ nói về các tạp chí và hội nghị quốc tế.

Có thể nói là do truyền thống, do bản chất của các khoa học, các ngành khoa học đều có “văn hóa ngành” của mình (theo cách gọi của GS Ngô Quang Hưng, Sate University of New York at Buffalo), trong đó có những thói quen, quy ước và ảnh hưởng khác nhau của các hội nghị khoa học trong ngành (Xem bài của Ngô Quang Hưng tại: <http://www.procul.org/blog/2005/03/24/cac-h%E1%BB%99i-ngh%E1%BB%8b-va-t%E1%BB%89-l%E1%BB%87-nh%E1%BA%ADn-bai/#postcomment>.)

Ở nhiều ngành khoa học, bài báo ở hội nghị khoa học được xem có giá trị thấp hơn rất nhiều so với bài ở tạp chí, thậm chí không đáng kể. Ở các ngành này, phần lớn các hội nghị chuyên ngành chỉ đòi hỏi người tham gia nộp một bản tóm tắt (abstract) hay tóm tắt mở rộng (extended abstract) chừng một hai trang. Hầu hết các bài (tóm tắt) gửi đến đều được nhận báo cáo tại hội nghị. Tại hội nghị chỉ có bản in các tóm tắt, và một số hội nghị sau đó sẽ chọn một số tác giả, yêu cầu nộp bài toàn văn để duyệt và in trong một tuyển tập.

Tin học (hay khoa học máy tính, KHMT) có “văn hóa ngành” khác. Các hội nghị KHMT đều yêu cầu người tham gia phải gửi bài toàn văn để tuyển chọn, với quy ước bài phải viết với một độ dài nào đó (như không quá 6000 từ, hoặc không quá 12 trang). Các hội nghị thường quy định nộp bài quãng 6 tháng trước ngày hội nghị để Ban chương trình có đủ thời gian tuyển chọn và in ấn. Mỗi bài gửi đến hội nghị thường được đánh giá bởi 3 đến 4 thành viên của Ban chương trình trong quãng thời gian 2 tháng. Gần đây nhờ Internet và các phần mềm giúp tuyển chọn bài, nhiều hội nghị đòi hỏi và cho phép tác giả trả lời những câu hỏi và ý kiến của các phản biện sau vòng duyệt đầu, và sau đó nhóm 3-4 phản biện của từng bài báo sẽ bàn bạc với nhau trước khi gửi quyết định của mình cho Trưởng Ban chương trình. Quá trình phản biện này thường chừng 2 tháng. Tuyển tập hội nghị, với các bài được chọn và sửa chữa theo góp ý của các phản biện, được in trước và có sẵn tại hội nghị. Uy tín của một hội nghị thường cũng đi cùng với uy tín của các thành viên Ban chương trình và nhà xuất bản. Tuyển tập của các hội nghị hàng đầu thường được in tại các nhà xuất bản hàng đầu như ACM, MIT, Springer, Morgan Kaufmann, ...

Có những xếp hạng chia các hội nghị ngành KHMT thành các nhóm rank 1, rank 2, rank 3, rank 4, ... hoặc không được xếp hạng, trong đó nhóm rank 1 có chất lượng cao nhất (như các hội nghị NIPS, IJCAI, ICML, ACM KDD, ... trong ngành Trí tuệ Nhân tạo) (có thể dễ dàng xem từ Google với cụm các từ “computer science conference ranking”, hoặc tương tự với các tạp chí bởi “computer science journal ranking”). Số bài được nhận báo cáo và đăng trong các tuyển tập thường quăng 25% số bài gửi đăng ký tham gia, trong đó nhiều hội nghị nhóm rank 1 có tỷ lệ bài được nhận khoảng 10-15%. Cần nói thêm rằng tỷ lệ nhận bài không đồng nhất với chất lượng hội nghị, vì chất lượng bài gửi đến các hội nghị ở các nhóm rank 1, rank 2, ... rất khác chất lượng bài gửi đến các hội nghị nhóm rank 4, hoặc các hội nghị chưa được xếp hạng. Nhiều cơ sở nghiên cứu ở các nước phát triển hằng năm có sẵn một danh sách các hội nghị “tốt”, chỉ tài trợ khi nhân viên gửi bài tham gia các hội nghị này.

Thật ra khó ai có thể biết rõ được cái “văn hóa” của một ngành nào đó nếu không làm việc trong ngành này. Theo quan sát của tôi, trong các ngành khoa học có truyền thống lâu năm như Toán học, Vật lý, Hóa học, Sinh học, Y học, ... các bài báo tạp chí được đánh giá cao hơn hẳn các bài báo ở tuyển tập của hội nghị. Trong khi đó các ngành khoa học non trẻ, như KHMT chẳng hạn, những bài báo ở các hội nghị chất lượng cao được đánh giá không thua kém gì bài ở các tạp chí tốt. Chẳng hạn, đại học Washington bang Seattle (một trong những đại học hàng đầu về KHMT ở Mỹ), khi xét tuyển chọn các chức vụ giáo sư, đánh giá một bài báo ở hội nghị AAAI tương đương một bài trong tạp chí rất tốt Journal of Artificial Intelligence Research.

Cái văn hóa ngành này còn thấy ở nhiều khía cạnh khác. Dễ thấy là các bài báo trong Vật lý, Hóa học, Sinh học, ... (những công trình liên quan đến công nghệ, thực nghiệm) thường do nhiều tác giả cùng làm, và bài ở nhiều tạp chí thường chỉ ngắn chừng vài ba trang, viết theo quy định của những khuôn mẫu nhất định (như các phần motivation, method, result, discussion, conclusion), trong khi đó bài ở các tạp chí về lý thuyết như toán học thường chỉ do một hai tác giả, và có thể dài nhiều chục trang. Chẳng hạn bài báo ở Annals of Math về kết quả dẫn đến giải thưởng của Viện Toán Clay của Ngô Bảo Châu và Gérard Laumon dài ... 99 trang.

Cũng nên nhìn nhận một thực tế là khoảng cách trong nhiều ngành khoa học của chúng ta so với thế giới đang từ khá cho đến rất xa, do vậy tuy đề cao những cần có các mục tiêu và tiêu chí về công bố quốc tế tại những nơi hợp với sức của từng ngành.

Góp ý này không có phần kết luận vì chỉ cung cấp thông tin cho mọi người.

Hồ Tú Bảo

Viện Khoa học và Công nghệ Tiên tiến Nhật bản

Các hội nghị và tỉ lệ nhận bài

Ngô Quang Hưng (trích từ Block Khoa Học Máy Tính, <http://www.procul.org/blog/>)

Trong khoa học máy tính, bài báo ở các hội nghị chuyên ngành danh tiếng đóng vai trò rất lớn trong sự nghiệp của các khoa học gia. Được nhận đăng bài ở một số hội nghị hàng đầu khó hơn ở nhiều journals có tiếng. (Xem thêm bài “tản mạn về mảnh bằng Ph.D” tôi viết vài năm trước.) Trong lý thuyết tính toán và giải thuật, ta có STOC, FOCS, SODA; trong cơ sở dữ liệu có SIGMOD, trong datamining có KDD, trong mạng máy tính có INFOCOM, SIGCOMM; vân vân. (Danh sách này không nhất thiết là đủ, nhưng khá đặc trưng cho các ví dụ này.)

Ở nhiều ngành khác (như toán, lý, xã hội học, ...) thì các journal papers giá trị hơn nhiều so với các conference papers. Ví dụ: các nhà toán học thường là không ghi các báo cáo ở hội nghị vào trong danh sách bài báo của họ.

Có vài nguyên do của sự “tréo ngoe” này trong ngành khoa học máy tính. Thứ nhất, KHMT phát triển cực nhanh trong vài thập niên gần đây, phần vì nó còn rất trẻ so với các ngành khác. Chờ khi công trình của mình được nhận đăng ở một journal (mất khoảng 1-2 năm) thì kết quả đó đã lỗi thời, thậm chí bản thân tác giả có khi cũng không thích thú gì lắm với nó nữa vì đã có những kết quả tốt hơn trong cùng thời gian. Vòng quay của các hội nghị tốn khoảng 6 tháng (từ khi nộp đến khi đi báo cáo). Thứ hai, đây cũng là vấn đề “văn hóa” của ngành. Thứ ba, hội nghị là một trong những phương tiện tốt nhất để mọi người làm quen, tìm hiểu nghiên cứu của nhau, tìm cơ hội hợp tác nghiên cứu, giới thiệu công trình của mình với thế giới các đồng nghiệp.

Các sinh viên, nhà nghiên cứu, giáo sư nào chưa có bài trong hội nghị lớn của ngành mình thì nói chung là kẻ ngoài cuộc, nghiên cứu không ai biết tới, và sẽ xa rời dòng chảy chính của các nghiên cứu trong ngành. Các bài báo này còn được dùng làm tiêu chí xét tenure, nhận giáo sư mới, thăng cấp giáo sư, vân vân.

Thế làm thế nào để biết là một hội nghị là “có giá” hơn các hội nghị khác? Dĩ nhiên người trong ngành sẽ biết (dù có thể hơi chủ quan nếu có hơn một hội nghị hàng đầu). Người ngoài ngành thì ... hỏi người trong ngành. Nếu không có ai để hỏi thì có thể tìm danh sách xếp hạng (ranking) các hội nghị (các danh sách loại này, dù là dựa trên chỉ số nào, cũng đều chủ quan và thiếu giá trị khoa học). Một cách nữa người ta cũng làm là nhìn vào tỉ lệ nhận bài của các hội nghị và danh sách các thành viên trong ủy ban chương trình kỹ thuật của hội nghị (technical program committee, hay TPC).

Thành viên TPC là những người sẽ đọc và quyết định bài nào được nhận, bài nào không. Ở các hội nghị lớn thì chất lượng TPC khá tương đồng. Như vậy chỉ số còn lại là tỉ lệ nhận bài (acceptance ratio). **Thế tỉ lệ nhận bài thấp có đồng nghĩa với giá trị cao của hội nghị không?** Graham Cormode, Artur Czumaj, và Muthu Muthukrishnan có [một bài rất khôi hài](#) (nhưng nghiêm túc) về các hội nghị trong khoa học máy tính và tỉ lệ nhận bài của chúng. Vấn đề chính họ muốn giải quyết là làm thế nào loại nhanh các bài báo tồi để các thành viên TPC đỡ mất thời gian.

Quay lại với câu hỏi trên. Câu trả lời dứt khoát là **không**. Đồng ý là có một tương quan nhất định giữa tỉ lệ nhận bài và giá trị hội nghị. Hội nghị nào (trong KHMT) có tỉ lệ nhận 50% hay nhiều hơn thì ta có thể tự tin kết luận là hội nghị thường thường bậc trung. Phần

còn lại thì rất khó nói. Những năm gần đây, MOBICOM nhận khoảng 8% đến 10%, INFOCOM nhận khoảng 16% đến 18%, còn STOC, FOCS, SODA nhận khoảng 25%-35%. Khó mà nói cái nào giá trị hơn cái nào trong các hội nghị trên, một phần vì chúng ở các nhánh khác nhau.

Lấy STOC và MOBICOM làm ví dụ. Đăng bài trong STOC rất khó, dù tỉ lệ nhận cao hơn MOBICOM khá nhiều. Một lý do là người ta thường không nộp các bài vớ vẩn vào STOC nữa. Ngoài ra chuyện này còn liên quan đến bản chất của ngành nghiên cứu. STOC là hội nghị về lý thuyết, kết quả tốt xấu khá rõ ràng. Ở các hội nghị đăng cả các bài báo thực nghiệm (simulation, experimentation) như mạng máy tính hay datamining thì kết quả không rõ ràng như thế, và sẽ có nhiều chỗ trống hơn cho các bài báo linh tinh. (Dù rằng các hội nghị danh tiếng thường chỉ đăng các bài có cơ sở lý thuyết vững chắc; phần simulation chỉ mang tính xác minh.)

Chuyện dài nhiều tập này xứng đáng vài posts nữa.