

Results of the First Survey of Practitioners of Evolutionary Computation

Gregory S. Hornby
University of California Santa Cruz
Mail Stop 269-3
Moffett Field, CA
hornby@email.arc.nasa.gov

Tina Yu
Computer Science Department
Memorial University of Newfoundland
St. John's, NL, Canada
tinayu@cs.mun.ca

Keywords

Survey, Evolutionary Computation, Technology Transfer, Real-World Applications

1. INTRODUCTION

The field of Evolutionary Computation (EC) has been around for several decades [2, 3], and in recent years there has been an explosion not only in the different types of biologically inspired algorithms, but also in the number of practitioners in the field. A critical part of this growth and development of the EC field has been the technology transfer of EC from academia to industry and the successful application of EC techniques to real-world problems. To assist in the continued technology transfer of EC techniques from academia to industry we conducted a survey of EC practitioners working in both academia and industry and in this article we summarize some of our findings.

The survey was conducted between March 1, 2005 and February 28, 2006 by posting 14 survey questions on the SIGEVO website. In particular, the survey asked about participants' background information, job information and obstacles encountered while applying EC techniques to industry jobs. Some of the main findings from our results are that: there has been an exponential growth in both EC graduates and practitioners; the main source for finding a job has been networking; while most respondents to our survey are in Europe, the most growth of EC in industry has been in North America; the main application areas of EC techniques are multi-objective optimization, classification, data mining and numerical optimization; and the biggest obstacle for the acceptance of EC techniques in industry is that it is poorly understood.

The survey we ran had three parts. First, it asked several questions about the participant's background. The second part of the survey had questions on their job information, and the third part was only for non-academic jobs and asked about EC acceptance and applications at that organization. In the rest of this article we present first the methodology of this survey and then our findings.

2. METHODOLOGY

The respondents to this survey were not randomly selected but were recruited through a variant of the snowball sampling strategy [4]. The recruiting methods include posting the survey announcement to various EC mailing lists (such as EC-Digest and genetic-programming), e-mailing the announcement to attendants of major EC conferences (such as GECCO-05, GPTP-05, EH-05) and advertising the survey at these conferences. Snowball sampling relies on referrals from the initial subjects to generate other subjects. Although snowball sampling may introduce bias into the study, it can be effective in reaching groups having common characteristics [1]. In our case, many EC practitioners are likely to subscribe to EC-related mailing lists and attend EC-related conferences, hence they can be reached by our recruiting approach. However, snowball sampling does not qualify as a random process. Consequently, the results from this survey cannot be generalized to the entire EC practitioner population, regardless of the number of responses received. Nevertheless, these results are still useful for gaining a preliminary picture of EC-practitioners in the world.

Over the one year time period in which the survey was taken, 324 responses were received, of which 305 had some EC relation, either through graduating with a degree specialized in EC or by using EC in one of their jobs. For the results of this survey, only the 305 responses which had an EC connection were used.

3. BACKGROUND INFORMATION

The first part of the survey asked participants to provide basic background information, such as gender, and to answer some questions about the most recent degree that they had received. We found that 71.1% of participants have a Ph.D. and the gender split is 87.5% male and 12.5% female. Looking into the geographic regions from which participants graduated from, we found that most participants graduated from Europe (46.2%), which is followed by North America (35.7%), Asia (12.5%), Oceania (2.6%), South America (2.0%), and Africa (0.7%). When we looked for yearly trends in these percentages based on graduation date, we found that they have remained fairly constant throughout the years.

One change that has occurred over the years is in the amount and specialization of graduates. The graduation rate has an exponential growth to it, starting with only a couple of people graduating a year from the 1960's up until the end of the 1980's, at which point the numbers increase dramatically and reach a peak of 36 graduates in 2004. For the first few decades, none of those who graduated in this time period had a degree specialized in EC. The first EC grad does not show up until 1991, and then starting in 1996 the majority

of graduates have an EC-specialized degree. This suggests that EC emerged as a field of its own sometime in the mid-1990s.

4. JOB SOURCES

After graduation, the next step is finding a job. By far the most common source of a job was networking, through which 35.1% of participants found a job. This was followed by *other* (25.7%), *supervisor* (17.2%), *postings at university department* (13.8%), *web* (4.7%), *campus career services center* (2.0%) and *mailing list* (0.7%). Looking for differences between those who took a job in academia versus those who took a non-academic job we found that *networking* was used more for finding a non-academic job (43.0%) than it was for finding one in academia (31.5%). In contrast, the reverse was true for *postings at the university department*: it was used by 16.5% of those who took an academic position but by only 8.0% of those who took a job in industry. Of those who selected *other*, 19 found their position through a listing in a journal or society magazine (such as the Communications of the ACM and IEEE), 13 found their job through an advertisement in the newspaper, 11 founded their own company, and 6 applied and received a research grant.

Further examination of the correlation between the job areas and the job-hunting methods found only a couple of patterns. One is that *postings at the university department* helped in finding jobs in Energy, Robotics and Government laboratories, but was of little use for the other job areas. Similarly, the *campus career center* had some success only in finding jobs in Government laboratories and Other. When job-finding methods are analyzed with the job regions, it shows some additional regional trends. *Networking* was used to find over half of the jobs in North America (as well as in Africa, Oceania and South America), but for less than a third of jobs in Europe, and for only 15% of jobs in Asia. In Europe, *supervisors* helped to find roughly a quarter of all jobs, and they were also helpful in Asia but were not very useful for finding jobs in North America. The *campus career center* was used by a small percentage of the respondents in Asia and North America, but was not used in any other geographic region.

5. JOB REGIONS

Looking into the distribution of jobs by geographic region, we found that most EC jobs have been in Europe (45%), followed by North America (37%), Asia (10%), Oceania (3%), South America (2%) and then Africa (2%). This geographic distribution of jobs closely matches the geographic distribution of graduates and suggests a strong correlation between where a graduate studied and where s/he worked. Also, the ratio of positions between the different geographic regions has been fairly constant over the years.

When the job positions are grouped by geological regions, analyzing the responses over the years reveals that the ratio between positions in industry and in academia has been fairly constant in recent years both in Europe (1:3) and in North America (2:3). In contrast, Asia has experienced a shift in its ratio from being predominantly in industry (100% non-academic in 1981) to being predominantly in academia (more than 75% academic in 2005). For the other geographic regions, the numbers of respondents was too small to give a meaningful interpretation.

Examining the movement of EC graduates for work reveals some interesting trends. First, none of the respondents who graduated with a degree specialized in EC from Africa or South America have left their regions for a job and only 12% of people who graduated

in Europe or North America ever move to a different region for work. In contrast, 44% of EC graduates in Asia and 40% of EC graduates in Oceania move at some point after graduation. Second, the direction of movement in Asia, Europe and North America is toward the West. Of those graduates who moved to a different geographic region for a job we found that: 62% of those graduating in Asia moved to Europe at some point, but only 25% ever moved to North America; 70% of those graduating in Europe moved to North America but only 20% moved to Asia; and 67% of those graduating in North America moved to Asia but only 17% moved to Europe for a job. Thirdly, for those people who moved from another region to North America, half moved for jobs in academia and half for jobs in industry, but for those participants who moved to a region other than North America, in all cases they went for academic positions.

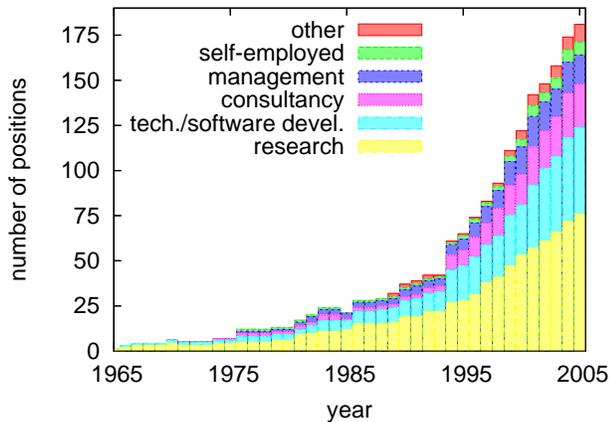
6. EC POSITIONS, PROBLEM TYPES AND APPLICATION AREAS

Once in a job, we are interested in what kind of position in their organization the respondent had, as well as whether or not EC was used and how it was applied. For determining trends by year each job entry, with its start and end years, was converted into yearly *positions*. That is, a job from 1997 to 2001 was separated into five positions, one each in 1997, 1998, 1999, 2000 and 2001. To limit participants such that they had at most one position in each year, jobs with overlapping years were modified so that the second job started in the year after the first job ended. For example, if a participant had a job from 1995 to 1998 followed by one from 1998 to 2001, the start year for the second job was changed from 1998 to 1999. Using this method, the 424 jobs that used EC techniques were mapped to 2955 EC-related positions.

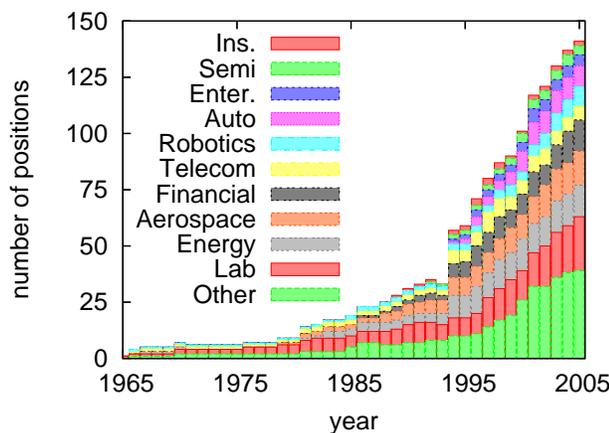
From our responses we found that there has been an exponential growth in positions in the field, starting with a single EC position in 1965 to just under 300 EC positions in 2005. Breaking this down into academic and non-academic positions, there has been a fairly steady proportion of just under a two-thirds of the positions in a given year being academic and just over a third being non-academic. Figure 1(a) is a breakdown of the type of position held for those not working in academia. This figure shows that most industrial EC positions are in research, with a significant number in technical/software development and consultancy.

Looking into the types of problems that respondents worked on we found the following: 40.3% do Multi-objective optimization (MOO), 38.4% do Numerical optimization, 38.0% do Classification, 37.7% do Other, 31.6% do Data mining, 21.2% do Open-ended design, 21.2% do Scheduling, 13.9% do Planning, and 10.1% do Satisfiability/TSP. These values do not add up to 100% because participants were able to make multiple selections for each job. For those responses that selected "Other", participants were able to enter a response in a text field. The most popular entries that were given are: optimization and design (24); modeling and simulation (17), EC theory (15); biology and bio-informatics (11); control (11); evolutionary robotics (6); artificial life (5) and neural networks (5). Many of these entries for "Other" fit under the given categories (eg. 'optimization and design' fits under Optimization and/or Open-ended design) with some of the other entries being an application area and not a problem type.

Comparing the distribution of problem-types worked on by academics to that of non-academics we found a significant difference (Table 1). In general, the percentage of academic positions that are working in a particular problem area is lower than that for non-



(a)



(b)

Figure 1: The figure contains boxplots of: (a) the type of industry position held; and, (b) the application area to which EC is applied.

academic positions. This means that academics tend to focus on fewer problem areas than those outside of academia. Specifically, those participants employed in academic positions average working on 2.24 problem areas whereas those in non-academic positions average working on 2.74 problem areas. Normalizing for this difference, *Scheduling* stands out as the one problem area which is significantly under-investigated by academics as compared to non-academics.

Different from the kind of problem being worked on (numerical optimization, scheduling, ...), is the industry to which this problem is being applied (automotive, insurance, ...). Figure 1(b) contains a histogram of EC industrial application areas by year.¹ The industry with the largest selection rate is *Other*, which was selected

¹Since each job was allowed to enter multiple application areas the total number of selected application areas can be greater than the number of positions.

Table 1: Percentage of respondents working in each problem area.

Area	Percentage (%) working in this area	
	Academia	Industry
MOO	38.8	45.9
Classification	38.0	46.6
Num. opt.	36.7	47.0
Other	39.4	34.4
Data mining	28.7	38.6
Scheduling	19.8	37.6
Open-ended design	24.7	24.2
Planning	13.5	21.3
Sat./TSP	10.1	15.6

in 37% of all jobs. The most common areas given by those who selected *Other* were: IT (13), consulting (12), biology/medicine related (e.g. Bioinformatics, biomedicine, pharmaceutical) (10), defense and military (7), and various types of engineering (civil, structural or manufacturing) (7). For non-academic jobs, the ways in which EC is reported to be most useful are: *design* (52.3%), *operations* (33.1%), *invention* (27.8%), *testing* (15.9%) and *other* (14.6%). Of the 31 responses for *other*, 10 were for optimization.

Next we looked into how application area varied by industry to see which combinations stand out (Table 2). Some specific combinations that we found are that those working in the automotive and robotics industries are interested in multi-objective and numerical optimization problems, while people working in the energy and entertainment industries are interested in multi-objective and classification problems. Finally, those working in insurance, telecommunications and the financial industries are predominantly interested in classification and data mining.

7. EC ACCEPTANCE IN INDUSTRY

Next we examined non-academic jobs to see what trends exist in the distribution and acceptance of EC in industry. Even though there is an exponential growth in the number of yearly EC positions, the ratios between the different levels of distribution and acceptance has remained fairly constant throughout the years. The acceptance rate has averaged: 41.3% well accepted; 19.8% accepted; 36.9% somewhat accepted; and 2.0% rejected. The distribution rate has averaged: 36.4% well distributed; 12.3% distributed; 25.3% somewhat distributed; and 26.0% isolated. That these ratios have remained fairly constant over the years does not mean that EC is not becoming more distributed and accepted in academic organizations – in fact, the growth in number of EC positions implies the opposite. What we cannot determine from our data is whether there is an increase in acceptance and distribution within an organization over time, and this is a question for a future survey.

We also analyzed EC acceptance by geographic region. The breakdown of acceptance in Asia, Europe and North America is as follows (well accepted, somewhat accepted, not well accepted, rejected): Asia (53%, 7%, 40%, 0%); Europe (41%, 35%, 20%, 4%); and North America (42%, 21%, 34%, 3%); We do not give a breakdown for the other geographic regions due to insufficient responses.

Important for increasing the acceptance and distribution of EC in industry is an understanding of the obstacles to its uptake. Based on our responses, we found the obstacles to be: *poorly understood*

Table 2: Percentages of job areas that involve work in different problem type.

Job Area	Tot #	Percentage working in this problem type.								
		Clsf	DM	MOO	NO	Design	Plan	S/TSP	Sched	Oth
Academic	27	30	33	48	56	19	7	11	15	30
Aerospace	20	50	45	60	55	30	15	20	40	25
Auto	10	40	60	80	80	20	40	20	50	50
Energy	20	70	50	65	50	15	45	10	50	20
Enter.	6	67	33	67	50	67	17	17	50	33
Financial	20	65	80	55	35	20	20	5	35	30
Gov. Lab	34	35	32	41	38	29	15	9	24	29
Insurance	5	100	80	60	20	0	20	0	20	0
Robotics	12	33	33	50	50	25	17	8	33	50
Semi-con	8	50	25	62	25	38	0	12	12	0
Tele-com	12	58	58	50	42	17	17	33	33	17
Other	56	48	38	54	36	20	21	9	32	41

(39.7%), *too ad hoc* (22.5%), *few successful applications to convince management* (21.2%), *commercial tools were unavailable or ineffective* (20.5%), *Other* (18.5%), *no proof of convergence* (14.6%), and *too hard to apply* (13.9%). In some ways, it is encouraging that the main obstacle is that EC is *poorly understood* because as more universities teach EC techniques, these methods should grow in familiarity and thereby gain wider acceptance in industry. Similarly, with a growth in familiarity of EC, companies may be less inclined to find it “ad hoc”. The third main obstacle is *the lack of successful applications*, is being addressed through Real-World Applications tracks at EC conferences and with the Human Competitive Competition held at GECCO since 2004. Finally, *lack of useful commercial tools* suggests a possible market niche for those wanting to achieve commercial success with creating EC software. Among the 27 responses for *Other*, the most common obstacles were: lack of experience/familiarity (9), and too slow or does not scale (4).

8. COMMENTS FOR FUTURE SURVEYS

Having conducted the first survey of practitioners of evolutionary computation we have some thoughts on changes that should be done for future surveys.

First off, to better understand EC education in universities it would be useful to ask for each degree received what the number of courses taken was in which EC techniques were covered. This would be beneficial for finding out how wide-spread EC techniques are being taught to non-EC specialists and also to find out if EC is being more widely included in course curriculum. Similarly, it would be useful to query people as to how many EC-specialized conferences they have attended in a given year, or the average number of such conferences they attended a year over the course of each job.

Second, in our questions on asking how well accepted/distributed EC is at a particular company, rather than having categories such as “Well accepted” to “Rejected or poorly accepted” for possible answers it would be more useful to ask for a numerical rating from 1 to 5, or 1 to 10 asking for the degree of acceptance. In this case 1 would be “Rejected” and the highest value would be “Well accepted.” Such a numerical system would allow for more fine-grained ranking of acceptance and would allow for numerical processing on how acceptance has changed. Also, it would be useful to ask for the level of EC acceptance at the start of a job and the level of EC acceptance and the end of the job (or its current level

of acceptance for jobs in which the respondent is still currently employed at). This would allow for analyzing whether there has been an increase in acceptance of EC at individual companies over time. Another question of use would be to ask for the size of the company or organization which the user is working at. It would be interesting to see if there are trends in the size of organization that uses EC, or in its growth in acceptance.

Finally, in addition to canonical evolutionary algorithms (such as genetic algorithms and evolutionary strategies) in recent years various other biologically-inspired computing algorithms such as ant colony optimization, artificial immune systems and particle swarm optimization have been developed. It would be useful to add a question asking respondents about which techniques they have used at each of their jobs so as to track their use and also learn what applications they are being used for.

9. CONCLUSION

Over the years, the use of EC techniques have grown from a few isolated practitioners into a genuine field with a large community. This first survey on EC practitioners has provided us with a preliminary picture of its development in the world. There has been an exponential growth in the number of EC practitioners and EC-specialized graduates, with the first graduates with EC-specialized degrees appearing in the mid 1990’s. After graduation, most survey participants found their jobs through networking or from their supervisors. Encouragingly, along with the growth in EC positions has been a growth in acceptance of EC techniques in industry, with the main obstacle to industry acceptance being that the technique is not well understood. EC has been applied to a wide variety of application areas and different problem domains, among which the most common problem areas are multi-objective optimization, classification, and numerical optimization. Although there are still challenges to the continued transfer of Evolutionary Computation to industry, we hope that the results of this survey will help.

10. ACKNOWLEDGMENTS

We would like to thank Gerardo Valencia for putting together the survey website, and also Erik Goodman, David Goldberg, John Koza and Una-May O’Reilly for their support in conducting this survey.

11. REFERENCES

- [1] R. Atkinson and J. Flint. Accessing hidden and hard to reach populations: Snowball research strategies. *Social Research Update*, 33, 2001.
- [2] K. A. De Jong. *Evolutionary Computation: A Unified Approach*. MIT Press, 2006.
- [3] D. B. Fogel, editor. *Evolutionary Computation: The Fossil Record*. IEEE Press, Piscataway, NJ, 1998.
- [4] W. P. Vogt. *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Science*. Sage, London, 1999.



Greg Hornby was a visiting researcher at Sony's Digital Creatures Laboratory from 1998 to 1999, received his Ph.D. in Computer Science from Brandeis University in 2002, and is currently a Project Scientist with U.C. Santa Cruz at NASA Ames Research Center.

His main area of research has been toward developing a science of generative representations and complexity measures for open-ended, computer-automated design. Among his accomplishments are the autonomous evolution of the dynamic gait for the consumer version of Sony's pet robot AIBO, and the evolution of the antenna for NASA's ST-5 Mission, which in 2006 became the first artificially-evolved hardware in space and the first evolved antenna to be used in an actual application. In addition, he developed the Age-Layered Population Structure (ALPS) Evolutionary Algorithm for reducing the problem of premature convergence.

Homepage: <http://ti.arc.nasa.gov/people/hornby>
E-mail: hornby@email.arc.nasa.gov



Tina received her Ph.D. from University of London, U.K in 1999. Between 1999 and 2005, she was with Chevron Technology Company. In 2005 she joined the Memorial University of Newfoundland as an Associate Professor.

Tina's recent research focuses on advancing the understanding and application of intelligent computing systems based on the modeling of biological processes, such as evolution, learning and cooperative behavior. This research incorporates theoretical analysis, statistical and empirical methods. The target results are new insights into system behavior, novel techniques to extend the current state-of-the-art and innovative real-world applications. As part of this focus she has co-edited 3 books in applying EC to real-world problems.

Homepage: <http://www.cs.mun.ca/tinayu>
E-mail: tinayu@cs.mun.ca