To be completed by all applicants

Confidential

Submitted: 11-24-2007, 18:58:35

# MIT Application for Graduate Admission 2008

*Please read instructions carefully before you complete this form. Please print clearly or type.*

1 Proposed date of entrance: ☐ February ☐ June ☒ September Year: 2008

2 Department Health Sciences and Technology Area of research or interdisciplinary program Bioinformatics and Integrative Genomics
(consult department listings in Book I)

3 Initial degree objective at MIT Ph.D. Final degree objective *(if different)* Ph.D.

4 Are you applying to more than one department? ☐ Yes ☒ No If yes, indicate departments _____

---

5 Full legal name Confidential
(last/family/surname) (first) (middle)

6 Date of birth Confidential
(month) (day) (year)

Former name (if any) _____

7 Female ☒ Male ☐

8 Ethnicity *(optional)*: **US citizens and permanent residents only**

☐ African-American/Black
☐ Afro-Caribbean
☐ African parentage
☐ Asian-American
☐ Caucasian/White
☐ Chicano or Mexican-American
☐ Puerto Rican
☐ Other Hispanic
☐ Native American _____ (Tribal affiliation)
☐ Other _____ (Please describe)

---

9 Reply address Confidential

10 Permanent ad

11 Daytime phon

12 Fax number _____
(country code) (area code/city code) (number)

---

13 City, state and country of birth Confidential United States of America Citizen of United States of America US Social Security # *(if an* Confidential

If a foreign citizen in the US, give date of entry _____ (month day year) Type of visa _____ I-20 ID Number _____

If an Exchange Visitor (J), give program number and name of sponsor _____

Permanent resident of _____ If a permanent resident (immigrant) of the US, give alien registration number _____

---

14 **Have you previously applied for admission to MIT?** ☐ Yes ☒ No MIT ID _____ (if assigned one as a former student)

If yes, what status? ☐ Freshman ☐ Transfer ☐ Special ☐ Graduate Date: _____ Department: _____

15 List all colleges and universities attended, major field, dates of attendance and name of degrees received or expected *(list most recent first)*:

| College/University | Location | Major field | Dates attended | Actual name of degree/diploma | Date degree awarded/expected |
|---|---|---|---|---|---|
| PRINCETON UNIV | PRINCETON, NJ | Computer Science | 09-2004 06-2008 | BA | 06-2008 |
| | | | | | |
| | | | | | |

16 Other graduate schools to which you are applying _____

For department use: ○ Admitted ○ Cond. Admitted Degree _____ Term _____ By _____ Date _____ ○ Not approved
○ Admitted ○ Cond. Admitted Degree _____ Term _____ By _____ Date _____ ○ Not approved

**MIT Application for Graduate Admission** *(continued from front)*

17 Entrance tests: GRE: Date taken or to be taken: 09-2007 Scores: verbal 800 quantitative 800 analytic 6

GRE Subject: Date taken or to be taken: _____ Scores: _____ Subject: _____

GMAT: Date taken or to be taken: _____ Scores: _____

TOEFL: Date taken or to be taken: _____ Scores: _____

18 List language of instruction in: primary school; English; secondary school English

university English; graduate school _____ Native language if other than English: _____

---

19 Names of three persons to whom you have given evaluation forms. *(Request those persons to return the completed forms to you in time to meet the appropriate deadline.)*

| name | title | institution/company |
|---|---|---|
| Dr. Olga Troyanskaya | Professor | Princeton University |
| Dr. Kevin Wayne | Professor | Princeton University |
| Dr. Brian Kernighan | Professor | Princeton University |

---

20 Please give the names, years of graduation, department affiliation, and relationships of any close relatives who have attended MIT:

21 Your honors, prizes, or major publications:

Sealfon et al., BMC Bioinformatics 7:443 [2006].

2007 Barry M. Goldwater Scholarship, Honorable Mention

22 Your extracurricular activities and accomplishments:

2006-present 2D Vegetarian Co-op, Webmaster

2006-present Princeton Art Museum Student Docent

23 Your teaching or professional experience including summer and term-time work. Give name of employer, dates, and nature of work:

Summer 2005-present Troyanskaya Laboratory, research in computational biology.

2007-present Undergraduate TA, COS 226 [Algorithms and Data Structures]

24 Other experience, including military, volunteer work, travel. Give dates and nature of work:

2007-present Peer Academic Advisor, Wilson Residential College.

Summer 2007 Orange Key Campus Tourguide

***International students*** *(non-immigrant visa holders) please complete the following:*

25 Your marital status: ☐ Single ☐ Married Number of children _____

If single, do you intend to marry before you come here? ☐ Yes ☐ No Will your spouse come with you? ☐ Yes ☐ No

26 Signature: Confidential Date: 11-24-2007

To be completed by all applicants except as noted

Confidential

Name: last/family first middle

Health Sciences and Technology DOC

MIT Department to which you are applying as a candidate for degree of

# Record of Courses Taken in Preparation for Graduate Study

*Please carefully read the instructions below before you complete this form.*
*Based on your transcript(s), please complete the following summary of your college and university classes.*

**General instructions:**

- Important: some departments do not require this form; other departments require only some of the fields to be completed—check the requirements for the department to which you are applying before completing this form. See pages 4–16 of this booklet.
- This form is **not** required for courses taken at MIT (except for the Department of Physics).
- If the department to which you are applying requires this form, a transcript will not be accepted as a substitute for this information.
- If your university system does not fit with the categories below—for example, if your courses do not have numbers or you did not receive course grades—leave those fields blank.
- Do not try to convert your university grading scale or GPA to MIT's scale. Indicate the grades/GPA as granted by your school and give us a brief explanation of your school's grading system.
- If the space provided is not sufficient, you may attach additional sheets.

**Why we ask you to complete this form:**

- We see transcripts from thousands of schools from all over the world. It is extremely helpful for us to review applicants' coursework and grades in a standard format.
- Transcripts show courses by semester or year. However, the best way for us to evaluate your preparation is to see your courses grouped by subject area, with the most relevant courses at the top.
- Transcripts do not list textbooks used; many departments find that information especially helpful.

**Grading System:**

Please describe the grading system(s) used at all colleges and universities you have attended. Explain the specific meaning of any numeric values, letter grades, and rankings. _____

_____

Cumulative GPA as listed on transcript (if available) _____

**COURSES MOST RELEVANT TO THE PROGRAM TO WHICH YOU ARE APPLYING**

**Please list below the college/university courses you have taken that are most relevant to the graduate program to which you are applying. Group courses by subject area, for example, group all math courses together and group all science courses together, etc. Use the next section, if you need more room.**

| Course no. (if applicable) | Course name | Principal textbook used (author and title) | Year in which course was taken | Official course grade (if applicable) |
|---|---|---|---|---|
| COS 423 | Theory of Algorithms | Corman, Leiserson, Rivest and Stein, Introduction to Algorithms | 2006-7 | A |
| COS 333 | Advanced Programming Techniques | Kernighan and Pike, The Practice of Programming | 2006-7 | A- |
| COS 402 | Artificial Intelligence | Russel and Norvig, AI: A modern approach | 2007-8 | |
| COS 981 | Junior Independent Work | | 2006-7 | A+ |
| COS 981 | Junior Independent Work | | 2006-7 | A |
| COS 226 | Algorithms and Data Structures | Robert Sedgewick, Algorithms in Java, 3rd Edition | 2005-6 | A+ |
| COS 217 | Introduction to Programming Systems | K.N. King, C Programming: A Modern Approach | 2005-6 | A |
| COS 341 | Discrete Mathematics | Kenneth H. Rosen, Discrete Mathematics and Its Applications, 6th edition | 2005-6 | A |
| MAT 306 | Graph Theory | Douglas B. West, Introduction to Graph Theory | 2005-6 | A |
| MAT 214 | Number Theory | Niven, Zuckerman, Montgomery, Introduction to the Theory of Numbers | 2004-5 | A |
| MathV1207 | Honors multivariate calculus & Linear Alg. I [At Columbia University, | Hubbard & Hubbard, Vector Calculus, Linear Algebra, and Differential Forms: A Unified | 2003-4 | A+ |
| MathV1208 | Honors multivariate calc. and linear alg II.[At Columbia University, | Hubbard and Hubbard Vector Calculus, Linear Algebra, and Differential Forms: A Unified | 2003-4 | A |
| ORF 309 | Probability and Stochastic Systems | Ross, Introduction to Probability Models | 2007-8 | |
| CHM 231 | An Integrated, Quantitative Introduction to the Natural Sciences | Tipler & Mosca, Physics for Scientists & Engineers. King, Java Programming. Zumdahl, | 2004-5 | A |

## ADDITIONAL COURSES

**Please list below other college/university courses you have taken. Group courses by subject area (science, math, engineering, humanities, social sciences, etc.). You may also use this section for any courses that you could not fit in the section above.**

| Course no. (if applicable) | Course name | Principal textbook used (author and title) | Year in which course was taken | Official course grade (if applicable) |
|---|---|---|---|---|
| CHM 232 | An Integrated, Quantitative Introduction to the Natural Sciences | Tipler & Mosca, Physics for Scientists & Engineers. King, Java Programming. Zumdahl, | 2004-5 | A |
| CHM 301 | Organic Chemistry I | Jones, Organic Chemistry | 2005-6 | A |
| CHM 302 | Organic Chemistry II | Jones, Organic Chemistry | 2005-6 | A |
| PHY 106 | Advanced Physics [Electromagnetism] | Purcell, Electricity and Magnetism | 2004-5 | A |
| ANT 206 | Human Evolution | McKee, Poirier, McGraw, Understanding Human Evolution | 2006-7 | A |
| | AP Biology | Campbell, Biology | 2002-3 | 5/5 |
| | AP Chemistry | Zuhmdahl, Chemical Principles | 2003-4 | 5/5 |
| | AP BC Calculus | Anton, Calculus | 2002-3 | 5/5 |
| ECO 101 | Introduction to Macroeconomics | Baumol and Blinder, Macroeconomics | 2004-5 | Pass |
| HEB 101 | Hebrew I | Brandeis University Textbook | 2005-6 | A |
| HEB 102 | Hebrew II | Brandeis University Textbook | 2005-6 | A |
| HUM 234 | East Asian Humanities: the Classical Tradition | Course packet | 2006-7 | Pass |
| WRI 123 | Introduction to Images: from Plato to Las Vegas | Course packet | 2004-5 | B+ |
| ENG 311 | Shakespeare II | Selected plays | 2004-5 | A |
| CWR 202 | Creative Writing Poetry | Norton Anthology of Modern and Contemporary Poetry | 2005-6 | Pass |
| CWR 202 | Creative Writing Poetry | Selected Poems | 2006-7 | Pass |
| CWR 301 | Advanced Creative Writing Poetry | Norton Anthology of Poetry | 2007-8 | |
| FRS 133 | 20th Century Poetry and Poets | Selected poems by Auden, Yeats, Eliot, Lowell | 2004-5 | A |
| HIS 383 | The United States Since 1920 | Irons, The Courage of Their Convictions. Boyer, Reagan as President, etc. | 2004-5 | A |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

THE BORDER OF THIS DOCUMENT IS PRINTED IN ORANGE INK

# PRINCETON UNIVERSITY

Con.

entered Princeton as a Freshman on 09/09/04.
rtment is Computer Science. This transcript prepared on 11/30/07.

| erm 2004-2005 (Freshman) | Grd | Hrs |
|---|---|---|
| Quantitative Intro/Natural Sciences | A | 4.0 |
| Quantitative Intro to Natural Sciences | A | 4.0 |
| 20th-Century Poems and Poets: Politics, | A | 4.0 |
| Writing Seminar | B+ | 4.0 |

| | Spring Term 2004-2005 (Freshman) | |
|---|---|---|
| ECO | 101 | Introduction to Macroeconomics |
| ENG | 311 | Shakespeare II |
| HIS | 383 | The United States since 1920 |
| MAT | 214 | Numbers, Equations, and Proofs |
| PHY | 106 | Advanced Physics (Electromagnetism) |

| erm 2005-2006 (Sophomore) | Grd | Hrs |
|---|---|---|
| Organic Chemistry I | A | 4.0 |
| Algorithms and Data Structures | A+ | 4.0 |
| Discrete Mathematics | A | 4.0 |
| Elementary Hebrew I | A | 4.0 |

| | Spring Term 2005-2006 (Sophomore) | |
|---|---|---|
| CHM | 302 | Organic Chemistry II |
| COS | 217 | Introduction to Programming Systems |
| CWR | 202 | Creative Writing (Poetry) |
| HEB | 102 | Elementary Hebrew II |
| MAT | 306 | Introduction to Graph Theory |

| Term 2006-2007 (Junior) | Grd | Hrs |
|---|---|---|
| Human Evolution | A | 4.0 |
| Advanced Programming Techniques | A- | 4.0 |
| Theory of Algorithms | A | 4.0 |
| Creative Writing (Poetry) | P | 4.0 |
| East Asian Humanities | P | 4.0 |
| Junior Independent Work | A | 4.0 |
| Junior Independent Work | A+ | 4.0 |

| | Fall Term 2007-2008 (Senior) | |
|---|---|---|
| COS | 402 | Artificial Intelligence |
| CWR | 301 | Advanced Creative Writing (Poetry) |
| ORF | 309 | Probability and Stochastic Systems |

its of Advanced Placement in Biology
its of Advanced Placement in Chemistry
its of Advanced Placement in French
its of Advanced Placement in Mathematics
arded THE SHAPIRO PRIZE FOR ACADEMIC EXCELLENCE
ıntary Withdrawal 12/15/06
ing 2006-2007 Readmitted
arded THE SHAPIRO PRIZE FOR ACADEMIC EXCELLENCE

RE

dmissions
518
:husetts Avenue
:, MA 02139

Polly Wintrey Griffin, Registrar

End of transcript

S OFFICIAL TRANSCRIPT IS PRINTED ON SECURITY PAPER AND DOES NOT REQUIRE A RAISED SE

# PRINCETON UNIVERSITY

## G SYMBOLS

**graduate courses (numbered below 500) and independent work**

Exceptional; significantly exceeds the highest expectations for undergraduate work
Outstanding; meets the highest standards for the assignment or course
Excellent; meets very high standards for the assignment or course
Very good; meets high standards for the assignment or course
Good; meets most of the standards for the assignment or course
More than adequate; shows some reasonable command of the material
Acceptable; meets basic standards for the assignment or course
Acceptable; meets some of the basic standards for the assignment or course
Acceptable, while falling short of meeting basic standards in several ways
Minimally acceptable; lowest passing grade
Failing; very poor performance
Grades of A+ through C- in courses taken on pass/D/fail basis (prior to 1988-89, earned grades of A+ through D were converted to P)
Satisfactory completion of required work in a course taken on an audit basis
Course not completed at end of term (late completion authorized)
Course successfully completed at another institution for Princeton credit
Course grades not reported by instructor
Student withdrew from the University during reading period

**ate courses (numbered 500 and above)**

exception of T and W, all of the foregoing grading symbols are used in graduate courses. The following symbols may also appear:

High Pass (used in some graduate courses in the School of Architecture)
Low Pass (used in some graduate courses in the School of Architecture)
No grade given in the course. Between 1948-49 and 1973-74, represented by N; from 1974-75, represented by *

## G POLICIES

004, the University faculty approved institutional grading expectations for undergraduate courses and independent work. These expectations posit a common grading for every department and program: A's (A+, A, A-) shall account for less than 35 percent of the grades given in undergraduate courses and less than 55 percent of the unior and senior independent work. Each department or program determines how best to meet these expectations. Beginning with fall term 2004-05, grades awarded tc iduates reflect the new expectations.

## E OF STUDY

***aduate students*** at Princeton enroll in a four-year course of study as candidates for the degree of Bachelor of Arts (A.B.) or the degree of Bachelor of Science in Engin The A.B. program consists of eight terms of full-time study to satisfy the requirement of 31 courses (30 courses for students matriculating before 2001). Beginning in the indidate for the A.B. degree undertakes a program of departmental concentration including course work, independent study in the junior year, a two-term senior thesis, a ental examination at the end of the senior year. The B.S.E. program consists of eight terms of full-time study to satisfy the requirement of 36 courses, which usually inclu /o terms of independent work. B.S.E. students pursue departmental concentrations beginning in the sophomore year. Prior to fall term 1974-75, an undergraduate's ental courses were indicated by a (D) preceding the course title. In addition to the departmental concentration, many students elect to pursue certificates in one or more s, nearly all of which are interdisciplinary. Notations regarding progress toward the degree: Suspension denotes an interruption resulting from disciplinary action. Requir val denotes an interruption resulting either from disciplinary action or from inadequate academic performance. Degree Withheld denotes seniors who have completed ei study, for whom the award of the degree has been delayed for disciplinary reasons. Failure to Qualify denotes seniors who have completed eight terms of study, but ha meet some significant University requirement, which postpones receipt of the Bachelor's degree.

***e students*** pursue full-time study toward the doctoral degree in the arts and sciences, engineering, architecture, and public affairs, and final professional master's degr ure, engineering, finance, Near Eastern studies, public affairs and public policy. To qualify for the Ph.D., a candidate spends at least one academic year in residence, pa ral examination, presents an acceptable dissertation, and passes the final public oral examination. Requirements for master's degrees vary by program. Graduate stude enrolled full-time and in residence hold regular student status as they pursue work toward the degree. Students registered *in absentia* are also enrolled full-time but are ipus in order to use materials and facilities not available in residence. The majority of post-generals graduate students will, in their last years of enrollment, take no cour ue full-time research on the dissertation. Students who come to the end of the normal period of enrollment without having completed all requirements for the Ph.D. hold egree candidacy continues for the first year and enrollment terminated/degree candidacy continues thereafter. These students are entitled to submit the dissertation at a

To be completed by all applicants

Confidential

Name: last/family first middle

Health Sciences and Technology

MIT Department to which you are applying

DOC

as a candidate for degree of

# Statement of Objectives

*Please read instructions carefully before you complete this form.*

Please give your reasons for wishing to do graduate work in the field you have chosen. Prepare your statement of objectives and goals in whatever form clearly presents your views. Include as far as you can, your particular interests, be they experimental, theoretical, or issue-oriented, and show how your background and MIT's programs support these interests. The statement could be much like a proposal for graduate studies, in the more specific context of your professional objectives. You should set forth the issues and problems you wish to address. Explain your longer-term professional goals. The Admissions Committee will welcome any factors you wish to bring to its attention concerning your academic and work experience to date.

- Type or print using black ink.
- Use reverse side if necessary or separate form.
- Keep a copy for your file. You may also wish to provide copies to your evaluators before they complete their forms.
- Return this form with the completed application to the appropriate MIT department (see pages 4–16).

My interest in biological research was stimulated by my research experience in a genomics laboratory that I worked in throughout high school. My freshman year at Princeton, I was lucky to have as my assigned academic advisor Professor Olga Troyanskaya, who works at the interface between biology and computer science. Although I had studied only one semester of computer science, she welcomed me into her lab at the end of my freshman year, and I have been there ever since. Through my computer science and mathematics coursework and research, I have come to appreciate that computational biology combines many of my academic passions. I greatly enjoy learning about the theoretical underpinnings of machine learning techniques and methods of data analysis. I love the satisfaction that comes from having built a working computer program. Most of all, I appreciate the ability to apply theoretical and computational techniques to answer fundamental biological questions relevant to human disease.

A problem of particular interest to me is the development of improved techniques to extract biological knowledge from large, diverse, and noisy datasets. In the Troyanskaya lab, I worked to create a visualization, exploration, and analysis system to help biologists effectively use Gene Ontology [GO]. GO is a controlled, hierarchical vocabulary for biological knowledge that is widely used as a trusted information source for genome-scale data analysis. My framework allows users to investigate large datasets in the context of the biological knowledge provided by GO. For example, a researcher who has analyzed microarray experiments and identified a cluster of genes that are co-regulated in a cancer study could use the system to identify known biological processes that are statistically enriched for these genes. To facilitate an understanding of the roles and relationships of the genes identified, the system allows the user to view a directed acyclic graph displaying a biological process of interest in the context of closely related known processes. I started this project as a participant in the Summer Programming Experience, a program for rising sophomores with little programming experience. After two years of work, the paper I wrote on GOLEM [Gene Ontology Local Exploration Map], was accepted by BMC Bioinformatics. Bio Med Central permanently designated it as "Highly Accessed," indicating that it was among the most frequently viewed articles published in the journal. Over 2000 researchers have accessed the paper. The software is also being integrated into other analysis systems for diverse biological data, and that work has been presented at the 2007 IEEE Parallel and Distributed Processing Symposium and described in the conference proceedings. As a result of the GOLEM publication, I was also contacted about a manuscript submitted to the journal Bioinformatics, and served as a peer reviewer.

Since last spring, I have been working to design a probabilistic Bayesian network to predict gene function in the malaria parasite. Malaria kills over a million children annually, and a better understanding of gene function in Plasmodium falciparum would help the development of drugs and vaccines to combat the parasite that causes the deadliest form of malaria in humans. Because over 60% of genes in P. falciparum are of unknown function, combining large scale experimental studies using machine learning techniques can further our understanding of the biology of this organism. By using Bayesian integration to combine heterogeneous datasets, I have generated a genome-scale functional interaction network for the malaria parasite. This integration comprises twenty-three diverse datasets, including microarray studies, mass spectrometry studies, a physical interaction study, and a number of datasets generated using in silico techniques. These datasets are first processed to identify the pairwise similarity of genes in each dataset, and then integrated using a Bayesian network trained and validated on a manually curated repository of known information on gene function in P. falciparum. The resulting functional interaction network includes around 14,000 novel high-confidence predicted interaction pairs. Among the highest-confidence interactions, I have detected a group of genes that may function to facilitate the process of erythrocyte invasion and that includes potential drug targets. In collaboration with the Llinas laboratory at Princeton University, which studies the biology of the malaria parasite, I plan to test some of my novel predictions in wet-lab

Confidential

11-24-2007

Date

Confidential

**Statement of Objectives (Cont.)**

experiments. I am excited about this project because of its clear relevance to human disease and its potential to provide insight into the biology of an organism about which much is currently unknown.

I am interested in teaching, as well as in conducting research. As an academic peer advisor this year, I have enjoyed the opportunity to advise freshmen interested in the sciences on their course choices. I have also made myself available through my residential college to students who want advice about computer science courses or the computer science department. In addition, I have worked as a laboratory TA for introductory computer science classes in the past, and I currently work as an undergraduate TA for a computer science course on algorithms.

Information from genome-scale experimental techniques is transforming our understanding of fundamental biological processes and accelerating insight into the mechanisms of human disease. I am particularly interested in applying machine learning and other computational techniques to the analysis, integration or interpretation of wet lab and in silico sources of biological information. In graduate school, I would like to conduct computational research on topics relating to human health, and to gain a solid background in relevant areas of biology, mathematics and computer science. Ultimately, I hope to pursue a career in computational medical research and teaching.

I would be very excited to pursue graduate studies in the HST program. I believe that the program's strong emphasis on interdisciplinary collaboration for medical research is a perfect match with my research interests. I am confident that the combination of research rotations, clinical experience, and coursework that the HST program offers will provide me with excellent training, and the Bioinformatics and Integrated Genomics track within the HST program is of particular interest to me. There are several laboratories and projects that potentially interest me, including the transcriptional regulatory networks research in the Bulyk laboratory, the application of computational techniques to the study of cancer in the Kohane laboratory, and the use of computational approaches to probe the relationships between sequence variation and health in the Sunyaev laboratory.

Confidential

**Use this space for additional information or comments, if needed.**

Confidential

Publications:
Sealfon, RSG, M.A. Hibbs, C. Huttenhower, C.L. Myers, O.G. Troyanskaya. "GOLEM: an interactive graph-based gene ontology navigation and analysis tool." BMC Bioinformatics 7:443 [2006].
-This paper has been accessed over 2000 times since publication and has received the permanent designation "highly accessed" from BioMed Central.

Wallace, G. , Hibbs, M., Dunham, M. Sealfon, R, Troyanskaya O, Li, K. "Scalable, Dynamic Analysis and Visualization for Genomic Datasets." Parallel and Distributed Processing Symposium, 2007.

Sealfon, Rachel Sima Gelernter. "The World in Black and White: Thomas Nast's Cartoons and the Fall of the Tweed Ring." The Concord Review. Volume 14,Winter 2004.
-This essay won the National Gilder Lehrman Prize in American History for the best essay in American history published in The Concord Review.

Reviewing Experience:
2006 Manuscript Reviewer for Bioinformatics

Honors and Awards:
2007 Elected to Phi Beta Kappa [one of 16 Princeton students in the class of 2008 elected at the beginning of senior year].
2007 Barry M. Goldwater Scholarship, Honorable Mention
2007 Anita Borg Memorial Scholarship Finalist
2006 Shapiro Award for Academic Excellence at Princeton University
2006 Outstanding Sophomore Award for Creative Writing
2005 Shapiro Award for Academic Excellence at Princeton University
2004 Maria Lo Frumento Award for achieving the highest GPA in the graduating class of 2004 at Hunter College High School.
2004 National Merit Scholar
2004 Siemens E. Westinghouse Competition Semifinalist
2004 The Elsbeth Kroeber Memorial Award "Presented in memory of a noted teacher and author by the New York Biology Teachers' Association for outstanding achievement in biology"

Employment:
2007-present Undergraduate TA, COS 226 [Algorithms and Data Structures]
Summer 2007 Bioinformatics research under Professor Olga Troyanskaya, Princeton University
Summer 2006 Bioinformatics research under Professor Olga Troyanskaya, Princeton University
2006 Lab Teaching assistant Computer Science 126, 127, 217, 226
Summer 2005 Bioinformatics research under Professor Olga Troyanskaya, Princeton University
2000-2004 Laboratory Research Intern, Mount Sinai School of Medicine

Activities
2007-present Peer Academic Advisor, Wilson Residential College, Princeton University
-Help advise a group of freshmen interested in studying science; answer students' questions about the computer science department and academic life at Princeton.
2006-present 2D Vegetarian Co-op, Webmistress
2007 Orange Key Tour Guide, Princeton University
2006-present Princeton Art Museum Student Docent

Princeton University
Lewis-Sigler Institute for Integrative Genomics
Princeton, NJ 08544
www.genomics.princeton.edu

Dear Admissions Committee-

It is my great pleasure to most highly recommend Confidential for graduate studies. I've known Con for over 3 years, first as her academic adviser and then as her research adviser for her summer research and junior independent work and currently senior thesis. To summarize my impression of Con – she is a true star: amazing, brilliant, thoughtful, creative, and independent -- an exceptional researcher in the making. I was thrilled when Con chose to major in Computer Science, and I am confident we will hear about Professor Con's great discoveries in applications of Computer Science techniques to biology in a few years time.

Why do I think this? Her accomplishments are clear from her academic record – straight A's and A+'s in some of the most challenging courses at Princeton and many academic awards including winning the Computing Research Association Undergraduate Award for 2008 (awarded to 3 undergraduates in the country), early induction into Phi Beta Kappa and Goldwater honorable mention, that all while performing challenging interdisciplinary research. What is perhaps even more important is her great intellectual curiosity complemented by her brilliance, technical excellence, and ability to focus on a specific problem at hand and identify original directions in that research area. During the summer after her first year at Princeton, Con took my group "by storm" by largely independently conducting a research project in bioinformatics with only half a semester of programming experience and no prior bioinformatics background. Over the next year and next summer, she completed the project by developing a novel system for Gene Ontology analysis, search, and exploration and wrote a paper about this, now published in *BMC Bioinformatics* (one of the main journals in the field). Her paper received stellar reviews and a lot of attention in the field, and her system is being used by biology researchers in many institutions. Her system is now being incorporated into a large genomic data management, search, and visualization framework that my lab is developing, and in that context Con's work has also been presented at the 2007 IEEE Parallel and Distributed Processing Symposium.

It is important to emphasize development of this visualization and analysis system was performed by Con largely independently – while I provided her with a problem to address, she identified specific directions in which to take the system and performed development, analysis, and evaluation of her system with minimal input from me. The publication that resulted from this work, of which Con is the first and lead author, was written by her as well, again with advice from me, but only on the level a senior graduate student might require. In summary, the research progress that Con has accomplished by the beginning of her junior year (working mostly during the summers) and the

scientific maturity she demonstrates is more than many graduate students at Princeton or other institutions do by the same time in their graduate careers.

[redacted]'s current work is a further development in the direction of developing novel computational systems for analysis of genomic data – she is developing a Bayesian framework for integration of diverse functional genomic data for the malaria parasite in a robust and accurate fashion. This is a very challenging problem – the data are highly heterogeneous (biologically and computationally) and extremely noisy, and learning has to be performed with very limited and biased gold standards. She already has a prototype system that performs better than any existing method and is working on improving it by extending the framework to integrate biological context. She plans to use this system to predict function for many of the unknown proteins in the malaria parasite, and, in collaboration with the laboratory of Manuel Llinas (a faculty member at the Molecular Biology department here), she will experimentally very the best of her predictions. This project further demonstrates [redacted]'s versatility and brilliance as a researcher – she is able to address both a challenging computer science problem and an exceptionally important biomedical problem. She is the leading (and only) computational researcher on this project – a position I generally reserve for senior graduate students or postdoctoral fellows – and she is making exceptionally fast and exciting progress. Given her preliminary results, I am confident this work will result in important contributions both to bioinformatics and malarial biology, and another lead author publication for [redacted]!

While [redacted] is clearly an exceptionally talented young researcher, she is also impressive in her interest and accomplishments outside of science, demonstrated by the diversity of her course interests and by the Sophomore Award for Creative Writing she received this year. In addition to this academic curiosity, or perhaps partly due to it, [redacted] is a wonderful collaborator, excellent presenter, and in general simply a great colleague to have. Even this early in her education, she is definitely not simply a student, but a true colleague: brilliant, creative, hard working, and an absolute delight to work with. I believe these qualities are in part responsible for her phenomenal success so early in her career in such an interdisciplinary field – one that requires knowledge and ability to speak the languages of computer science, statistics, and biology.

To summarize, [redacted] is *the best* undergraduate I have encountered in my 8 years at Princeton and Stanford, and can be easily compared to the strongest first and second year graduate students at these institutions. I have no question she will excel in graduate school, and I would love for her to join my lab.

Please do not hesitate to contact me if you have any further questions.

Olga Troyanskaya
Olga Troy
Assistant Professor
Department of Computer Science & Lewis-Sigler Institute for Integrative Genomics
Princeton University

# MIT Evaluation for Graduate Admission

*Please read instructions carefully before you complete this form.*

**Return to:**
Department of _____
Massachusetts Institute of Technology
77 Massachusetts Avenue, Room _____
Cambridge, MA 02139-4307

**Part 1 To be completed by all applicants**

Please type or print using black ink.
Important: In the upper right of this form, fill in the return address department name and room number (as indicated on pages 4–16 of this booklet).

Name: Confidential _____ _____
last/family first middle

Applying for admission in the department of Health Sciences and Technology

for the ultimate degree of Ph.D. area of research BIG

for the term beginning in September 2008 _____ email Confidential

Under the Family Educational Rights and Privacy Act of 1974, a student enrolled at MIT has access to his or her education records. We intend to comply with both the letter and the spirit of this law, while still allowing the student the option of waiving the right to access. If you wish to waive the right to examine this evaluation at a later date, please sign here.

Applicant's signature: Confidential _____
date

**Part 2 To be completed by evaluator**

An application for admission to MIT requires evaluations from three teachers or people capable of judging the professional and academic promise of the applicant. A separate letter of evaluation may be attached to this form if necessary.

Please return in time for her/him to meet the following deadlines: January 15 for June or September admissions for applicants except as follows: December 15 for Aeronautics and Astronautics, Architecture, Biology, Brain and Cognitive Science, Chemistry, Electrical Engineering and Computer Science, Health Sciences and Technology, Leaders for Manufacturing (see http://lfm.mit.edu for additional details), Mechanical Engineering, Media Arts and Sciences; December 15 for Operations Research; December 31 for Biological/Engineering and Political Science; January 1 for Physics and Science, Technology and Society; January 2 for Chemical Engineering, Civil and Environmental Engineering, Economics, Linguistics and Philosophy, Mathematics; January 3 for Urban Studies and Planning; January 5 for Earth, Atmospheric and Planetary Sciences; January 7 for Nuclear Science and Engineering; January 10 for Computation for Design and Optimization, Engineering Systems Division and Technology and Policy Program; and February 15 for Center for Real Estate. November 1 is the deadline for the February term. January 12 and April 6 are the deadlines for Round 1 and Round 2, respectively, for the Master of Engineering in Logistics (MLOG) Program. For the Evaluation Form, please go to http://www.mit.edu/mlog/. For the Systems Design and Management Evaluation Form, please go to http://sdm.mit.edu.

Evaluator's name: Kevin Wayne Title Senior Lecturer

Address 52 Cedar Lane Princeton NJ 08540 USA

email wayne@cs.princeton.edu Date 11-28-2007

School or company Princeton University Telephone number 609 258-4455

In what capacity do you know the applicant? instructor, supervisor

How long have you known the applicant? 3 1/2 years

How does this applicant compare with his or her peer group in academic ability?

☒ Truly exceptional *equivalent to the very best you have known – a person who, in your experience, appears only every few years*
☐ Outstanding *comparable to the best student in a current class*
☐ Well above average *top 25%*
☐ Above average *demonstrated high ability*
☐ Average *able to complete work to the Ph.D.*
☐ Below average *lower 50%*
☐ Inadequate opportunity to observe

In your opinion, how would this student compare to other students in the graduate program at MIT?

☒ Truly exceptional ☐ Outstanding ☐ Well above average ☐ Above average ☐ Average ☐ Below average ☐ Inadequate opportunity to observe

Please give the applicant's relative standing in your department (e.g., 7th in 89) 1 of 40

*(continue on reverse side)*

Personal evaluation of the applicant. What particularly qualifies this student for study at MIT? Information about accomplishments in research or independent projects will be particularly helpful. If you know of other students who have entered MIT from your institution, a comparison will be especially valuable. If you have any reason to believe that the applicant should not be considered, please explain.

See attached pdf.

If the applicant's first language is not English, please evaluate her/his proficiency to read, write, and speak English.

Evaluator's signature Kevin Wayne 11-28-2007

Evaluators: Please feel free to add information about your own educational and professional background if you feel that such information will enhance our understanding of your evaluation.

*Please seal and sign the envelope.*

November 28, 2007

Dear Graduate Admissions Committee,

It is my pleasure to most strongly and fervidly recommend Con for admission into your Ph.D. program. I have gotten to know Con over the past three and a half years, and have been blown away by her abilities. She is one of the most exceptional and promising students I have known in my 9 years at Princeton (out of over 1000). I have no doubt that she will continue to be a star student here, and believe that she has the potential to grow into a superstar researcher as she continues her studies at Princeton and beyond.

In Fall 2004 I taught the computer science portion of an extremely intensive and ambitious science curriculum that integrates biology, chemistry, computer science, and physics into a two-semester double-course for freshman. Rachel entered the course with absolutely no background in computer science or programming. By the end of the class, she impressed me so much that I wanted to convince her to become a computer science major!

In Fall 2005, I was delighted to have Con in COS 226, an intermediate level data structures and algorithms course. Despite my very high expectations, Con continued to impress, surprise, and amaze me with her algorithmic and mathematical talents. She was the outstanding student in the class, earning the only A+. I could hardly wait to see what she accomplishes when she gets beyond coursework and sinks her teeth into research problems. I also got to know Con on a personal level in office hours. In addition to her intellectual gifts and thirst for knowledge, she is among the nicest, most modest, and most respectful people I have ever met. It has been a true privilege and joy to have her as a student.

In Summer 2006, I directed a summer program called *Summer Programming Experiences* that provide talented freshmen an opportunity to carry out independent work, under the supervision of a faculty member or graduate student. I recruited Con to apply and was delighted when she decided to participate. Con worked in the Troyanskaya lab and developed a gene ontology browsing tool for analyzing genomic data and structuring of biological knowledge. This project involved a number of data visualization

components, including depth-based layout, graph visualization, annotations, statistical analysis and searching functionality. The project was such a success that she continued to work on the project after the program ended, and published the result (as the lead author) in *BMC Bioinformatics*. The tool is now widely used within the bioinformatics community. This is an extraordinary accomplishment, beyond even my wildest expectations. As a freshman, she was already performing at the level of a 2nd or 3rd year graduate student. From this experience, I can see that Con is an extremely hard-working, driven, and brilliant scientist who is destined to accomplish great things.

This semester, I recruited Con to serve as an undergraduate teaching assistant in an introductory algorithms and data structures course. She was one of two students we hand-picked for the newly created job. Although she has only been working for two months, she is already doing a better job than any of the 50+ graduate TAs I've had over the past 10 years! Her clarity of explanations, attention to detail, and mastery of the material far exceeds anything I would expect from a second year graduate student, let alone an undergraduate.

I believe that Con is quite capable of doing anything she puts her mind to, whether it be in computer science or bioinformatics. I hope that your Ph.D. program provides her with an opportunity to do exactly that. I can't think of anyone more qualified or more worthy (and this includes a very distinguished group of Marshall, Rhodes, NSF, Goldwater, and CRA scholarship winners whom I've been fortunate enough to teach and advise).

Sincerely,

Kevin Wayne
*Phillip Y. Goldman '86 Senior Lecturer*
*Department of Computer Science*
*Princeton University*

# MIT Evaluation for Graduate Admission

*Please read instructions carefully before you complete this form.*

**Return to:**
Department of _____
Massachusetts Institute of Technology
77 Massachusetts Avenue, Room _____
Cambridge, MA 02139-4307

**Part 1** To be completed by **all applicants**

Please type or print using black ink.
Important: In the upper right of this form, fill in the return address department name and room number (as indicated on pages 4–16 of this booklet).

Name: Confidential
last/family first middle

Applying for admission in the department of Health Sciences and Technology

for the ultimate degree of Ph.D. area of research BIG

for the term beginning in September 2008 email Confidential

Under the Family Educational Rights and Privacy Act of 1974, a student enrolled at MIT has access to his or her education records. We intend to comply with both the letter and the spirit of this law, while still allowing the student the option of waiving the right to access. If you wish to waive the right to examine this evaluation at a later date, please sign here.

Applicant's signature: Confidential
date

**Part 2** To be completed by **evaluator**

An application for admission to MIT requires evaluations from three teachers or people capable of judging the professional and academic promise of the applicant. A separate letter of evaluation may be attached to this form if necessary.

Please return in time for her/him to meet the following deadlines: January 15 for June or September admissions for applicants except as follows: December 15 for Aeronautics and Astronautics, Architecture, Biology, Brain and Cognitive Science, Chemistry, Electrical Engineering and Computer Science, Health Sciences and Technology, Leaders for Manufacturing (see http://lfm.mit.edu for additional details), Mechanical Engineering, Media Arts and Sciences; December 15 for Operations Research; December 31 for Biological/Engineering and Political Science; January 1 for Physics and Science, Technology and Society; January 2 for Chemical Engineering, Civil and Environmental Engineering, Economics, Linguistics and Philosophy, Mathematics; January 3 for Urban Studies and Planning; January 5 for Earth, Atmospheric and Planetary Sciences; January 7 for Nuclear Science and Engineering; January 10 for Computation for Design and Optimization, Engineering Systems Division and Technology and Policy Program; and February 15 for Center for Real Estate. November 1 is the deadline for the February term. January 12 and April 6 are the deadlines for Round 1 and Round 2, respectively, for the Master of Engineering in Logistics (MLOG) Program. For the Evaluation Form, please go to http://www.mit.edu/mlog/. For the Systems Design and Management Evaluation Form, please go to http://sdm.mit.edu.

Evaluator's name: Brian Kernighan Title Professor

Address 35 Olden St Princeton NJ 08540 USA

email bwk@cs.princeton.edu Date 11-19-2007

School or company Princeton University Telephone number 609-258-2089

In what capacity do you know the applicant? professor in 1 class; class advisor

How long have you known the applicant? 1 year

How does this applicant compare with his or her peer group in academic ability?

- ☐ Truly exceptional — *equivalent to the very best you have known – a person who, in your experience, appears only every few years*
- ☒ Outstanding — *comparable to the best student in a current class*
- ☐ Well above average — *top 25%*
- ☐ Above average — *demonstrated high ability*
- ☐ Average — *able to complete work to the Ph.D.*
- ☐ Below average — *lower 50%*
- ☐ Inadequate opportunity to observe

In your opinion, how would this student compare to other students in the graduate program at MIT?

- ☐ Truly exceptional
- ☐ Outstanding
- ☒ Well above average
- ☐ Above average
- ☐ Average
- ☐ Below average
- ☐ Inadequate opportunity to observe

Please give the applicant's relative standing in your department (e.g., 7th in 89) top 3-4 of 40

*(continue on reverse side)*

Personal evaluation of the applicant. What particularly qualifies this student for study at MIT? Information about accomplishments in research or independent projects will be particularly helpful. If you know of other students who have entered MIT from your institution, a comparison will be especially valuable. If you have any reason to believe that the applicant should not be considered, please explain.

please see attached letter

If the applicant's first language is not English, please evaluate her/his proficiency to read, write, and speak English.

Evaluator's signature Brian Kernighan 11-19-2007

Evaluators: Please feel free to add information about your own educational and professional background if you feel that such information will enhance our understanding of your evaluation.

*Please seal and sign the envelope.*

November 19, 2007

Letter of Recommendation for Confidential

Con was a student in my Advanced Programming Techniques course (COS 333) in Spring 2007. COS 333 is quite demanding, and covers a very wide range of programming languages, tools, and techniques, including scripting languages like Perl, object-oriented programming in C++ and Java, a taste of relational databases and network programming, and graphical interface tools like Swing and Visual Basic.

In addition to individual programming assignments, there is also a group design and implementation project that keeps the students busy for two thirds of the term and accounts for two thirds of the course grade. The project includes not just design and implementation but weekly meetings with a teaching assistant to keep track of progress, a public presentation and demo of the result at the end of the semester, and an extensive written report. In addition, all code is submitted for examination, and the TAs and I all experiment with the systems.

The general structure for the project is a three-tier system, with a user interface, a database, and some logic in the middle. Within that loose framework, the students invent their own projects, pursuing whatever interests them.

Con 's group of two juniors and a sophomore had a really neat idea for a project: create a site that analogous to the existing student course guide that would provide information for students intending to pursue independent work. This is quite important at Princeton, where almost everyone is required to do two junior projects and a senior thesis. Their system was meant to give information on previous theses, the research interests of faculty and what they had supervised, and provide a way for students to provide anonymous comments on how well professors served as advisors.

The system was written in largely in PHP, about 2000 lines, with a MySQL database to maintain the information. There was a fair amount of ancillary code in other languages to extract information from the university's somewhat erratic systems; all of this was neatly integrated into database tables by an automated process.

The system was very well done, for the most part visually attractive, and provided easy and intuitive access to some fascinating information that

could be collated and searched on a variety of axes, mostly through a very dynamic Ajax interface. I found myself wasting far too much time poking around to see who had written what and who supervised it. At the end of the semester, the group gave a very polished presentation, clearly the result of serious rehearsal, and the final report was very well written and (most unusual) almost entirely free of errors. (In hindsight, I expect that this high quality writing was largely due to Con, who has a separate life as a fine writer.) This project is one of the ones that have some prospect of being genuinely useful with some more effort, if one were able to fight the uphill battle of getting commonality among the various data sources. It's probably too late now that Con and her colleagues have moved on to other work, but I would really like to see the system maintained.

Academically, Con is right at the top of our population, in a tight clump of three very strong students. In my own class she was not the strongest programmer of that specific trio, but she's the one most likely to win a major award for her scientific research -- she has all the right skills and she's interested in an area with great potential. She's been doing remarkable work in computational biology almost since she arrived here, largely with my colleague Olga Troyanskaya (who has an unerring eye for outstanding young people). In my class, Con was a bit quiet -- indeed I wish that she had said more in my class, since everything she said was right on target -- but she's a genuinely nice person, one who would make a great colleague. She's already been a role model for other women students both in residences and through her work as an undergrad lab assistant for CS courses. She's marked for great things, and I recommend her in the strongest terms.

Brian Kernighan
Professor
Department of Computer Science
Princeton University