What flies can teach us about searching the web

Educator guide

PAPER DETAILS

Original title: A neural algorithm for a fundamental computing problem

Reference: Vol. 358, Issue 6364, pp. 793-796

Authors: Sanjoy Dasgupta, Charles F. Stevens, Saket Navlakha

Issue name: *Science*

Original publication date: 10 Nov 2017

DOI: 10.1126/science.aam9868

Annotator(s): Iris Stone

TABLE OF CONTENTS

DISCUSSION QUESTIONS

	LEARNING STANDARDS
1. What is the purpose of a similarity search? Can you think of some ways that this type of algorithm appears in your everyday life?	**SEP6** **ETS2.B** **SP7**
2. Why do you think locality-sensitive hashing is so often performed in similarity searches? What aspect(s) of the algorithm make(s) it useful for solving a similarity search problem?	**SEP5** **ETS1.C** **Systems and System Models**
3. How do the structure and placement of cells in the fly olfactory circuit (e.g. ORNs, PNs, KCs, and APLs) help them perform their role as part of an algorithm? *(Note: This might require students to do some additional research on what cells look like and where they're located in the fly olfactory circuit)*	**LS1.A** **Structure and Function** **EK3.E.2**
4. What can you conclude from this study? What might the authors' next steps be in terms of designing a follow-up experiment?	**SEP1** **RST.11-12.2** **SP3** **Nature of Science**
5. In the abstract, the authors write that, the "fly algorithm, however, uses three computation strategies that depart from traditional approaches." What three strategies are the authors referring to?	**ETS1.C** **Structure and Function** **RST.9-10.2**
6. Can you think of any considerations that the authors overlooked? If you had to offer suggestions for how they could improve their study, what would you say?	**SEP1** **ETS1.C** **SP3** **RST.11-12.6** **RST.11-12.8**

ACTIVITIES FOR INTERACTIVE ENGAGEMENT

	LEARNING STANDARDS
Writing an abstract Students write a new abstract for the article at a grade-appropriate reading level.	**RST.9-10.2** **RST.11-12.2** **Nature of Science**
Locating this study in the larger field Students use the annotated list of references to explain how this research builds on the published work of at least one other independent group of scientists. Students will evaluate whether data from this research supports or contradicts previous conclusions and reflect on the statement that scientific knowledge is a "community effort."	**RST.9-10.8** **RST.11-12.8** **SP3** **Nature of Science**
Science in the news Students explore news stories in the Related Resources tab and evaluate the stories for tone, accuracy, missing information, etc. They may then write their own news stories on the article.	**RST.9-10.9** **RST.11-12.2** **RST.11-12.8** **RST.11-12.9**
Creating a model Students draw models/schematics of the LSH and fly algorithms and label them with specific terms, including the values of variables used in the authors' calculations.	**SEP2** **SEP5** **LS1.A** **Systems and System Models** **SP1**
Biomimicry and human design Given that the authors propose that computer search algorithms could be improved by following the lead of fruit fly olfactory systems, students will design other opportunities for human-designed systems to be improved using nature as a guide. Students can also examine other examples of human-designed systems already in use that mimic naturally-occurring systems. *(Note: This can potentially lead to an exploration of an article in* Science *about robots mimicking fruit fly flight*: http://science.sciencemag.org/content/361/6407/1073*)*	**ETS2.B** **Structure and Function** **SP7** **Nature of Science**
Results and conclusions Students diagram each of the experiments presented in the study (divided up by figure, if appropriate). They then consider the results depicted in each figure, and how these results support the conclusions of the study.	**SEP2** **Systems and System Models** **RST.11-12.8** **SP1** **Nature of Science**
The next steps Students design a follow-on experiment to this study that either addresses flaws or unanswered questions in the research at hand or builds on it to explore a new question.	**SEP1** **RST.11-12.5** **RST.9-10.8** **SP7** **Nature of Science**

ARTICLE OVERVIEW

Article summary (recommended for educator use only)

Authors use the fruit fly olfactory neural circuit to address a fundamental task in computer science and data retrieval systems. This task, called similarity search, involves identifying similar items (e.g. images, words, or websites) in a database (e.g. the internet). Computer scientists typically conduct similarity searches using a type of algorithm called locality-sensitive hashing, or LSH. Interestingly, the authors discover that the algorithm that the fly olfactory circuit uses to identify odors has many similarities to LSH, save for a few important distinctions. When the authors modified a traditional LSH algorithm to incorporate fly-specific strategies, they were able to markedly improve the algorithm's ability to perform similar searches.

Importance of this research

The authors' goal in this study was to find a better way to solve a common problem in large-scale information retrieval systems (such as online search engines). To do so, they hypothesized that the brain might naturally use computational methods that are superior to what we currently use in computer science. Specifically, they decided to focus on a part of the brain in the fruit fly known as the olfactory circuit. This decision was motivated by previous fruit fly research that elucidated details about how the fly brain identifies odors; what types of cells are involved in the computations; and how those computations store, organize, and transmit information. The results of this study confirm the authors' hypothesis and suggest that the brain may provide a computational model for many different types of problems. Future research will likely focus on (1) how to apply these results to other computational tasks and (2) what other areas of the brain might inspire their own solutions to important computing problems.

Experimental methods

- Empirical comparison of algorithms using computational modeling
- Testing of algorithmic functionality using benchmark data sets
- Quantification of algorithm performance using mean average precision
- Designing a locality-sensitive hashing algorithm using proofs from the Johnson-Lindenstrauss lemma

Conclusions

- The fruit fly olfactory circuit identifies odors using a type of computer science algorithm known as locality-sensitive hashing (LSH).
- There are three aspects of the fly's algorithm that differ from traditional LSH approaches.
- Applying these three fly-specific strategies to a traditional LSH algorithm improves its ability to solve a common computational problem.

LEARNING STANDARDS ALIGNMENT

The following tables provide an overview of the learning standards covered by this article, including the A Framework for K-12 Science Education (Framework), Common Core State Standards English Language Arts-Literacy (CCSS ELA), Common Core State Standards Statistics and Probability (CCSS HSS), AP Science Practices, and Vision and Change for Undergraduate Education. Where applicable, activities and information will be marked with specific standards to which they are linked.

A Framework for K-12 Science Education		
Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Asking Questions and Defining Problems(SEP1) Evaluate a question to determine if it is testable and relevant. Ask and/or evaluation questions that challenge the premise(s) of an argument, the interpretation of a data set, or the suitability of a design. **Developing and Using Models (SEP2)** Develop, revise, and/or use a model based on evidence to illustrate or predict the relationships between systems or between components of a system. **Using Mathematics and Computational Thinking (SEP5)** Use mathematical, computational, and/or algorithmic representations of phenomena or design solutions to describe and/or support claims and/or explanations.	**LS1.A: Structure and Function** Systems of specialized cells within organisms help them perform the essential functions of life. Feedback mechanisms maintain a living system's internal conditions within certain limits and mediate behaviors, allowing it to remain alive and functional even as external conditions change within some range. Feedback mechanisms can encourage (through positive feedback) or discourage (negative feedback) what is going on inside the living system. **ETS1.C: Optimizing the Design Solution** Criteria may need to be broken down into simpler ones that can be approached systematically, and decisions about the priority of certain criterial over others (trade-offs) may be needed. **ETS2.B: Influence of Engineering, Technology, and Science on Society in the Natural World** New technologies can have deep impacts on society and the environment, including some that were not anticipated. Analysis of costs and benefits is a critical aspect of decision about technology.	**Systems and System Models** Models (e.g. physical, mathematical, computer models) can be used to simulate systems and interactions —including energy, matter, and information flows —within and between systems at different scales. **Structure and Function** The functions and properties of natural and designed objects and systems can be inferred from their overall structure, the way their components are shaped and used, and the molecular substructures of its various materials.

Common Core State Standards English Language Arts-Literacy		
Key Ideas and Details	Craft and Structure	Integration of Knowledge and Ideas
RST.9-10.1 Cite specific textual evidence to support analysis of science and technical texts, attending to the precise details of explanations or descriptions. **RST.9-10.2** Determine the central ideas or conclusions of a text; trace the text's explanation or depiction of a complex process, phenomenon, or concept; provide an accurate summary of the text. **RST.11-12.1** Cite specific textual evidence to support analysis of science and technical texts, attending to important distinctions the author makes and to any gaps or inconsistencies in the account. **RST.11-12.2** Determine the central ideas or conclusions of a text; summarize complex concepts, processes, or information presented in a text by paraphrasing them in simpler but still accurate terms.	**RST.9-10.4** Determine the meaning of symbols, key terms, and other domain-specific words and phrases as they are used in a specific scientific or technical context relevant to grades 9-10 texts and topics. **RST.9-10.5** Analyze the structure of the relationships among concepts in a text, including relationships among key terms (e.g., force, friction, reaction force, energy). **RST.9-10.6** Analyze the author's purpose in providing an explanation, describing a procedure, or discussing an experiment in a text, defining the question the author seeks to address. **RST.11-12.4** Determine the meaning of symbols, key terms, and other domain-specific words and phrases as they are used in a specific scientific or technical context relevant to grades 11-12 texts and topics. **RST.11-12.5** Analyze how the text structures information or ideas into categories or hierarchies, demonstrating understanding of the information or ideas. **RST.11-12.6** Analyze the author's purpose in providing an explanation, describing a procedure, or discussing an experiment in a text, identifying important issues that remain unresolved.	**RST.9-10.8** Assess the extent to which the reasoning and evidence in a text support the author's claim or a recommendation for solving a scientific or technical problem. **RST.9-10.9** Compare and contrast findings presented in a text to those from other sources (including their own experiments), noting when the findings support or contradict previous explanations or accounts. **RST.11-12.8** Evaluate the hypotheses, data, analyses, and conclusions in a science or technical text, verifying the data when possible and corroborating or challenging conclusions with other sources of information. **RST.11-12.9** Synthesize information from a range of sources (e.g., texts, experiments, simulations) into a coherent understanding of a process, phenomenon, or concept, resolving conflicting information when possible.

AP Science Standards	
AP Science Practices	AP Biology Content Standards
Science Practice 1 (SP1) The student can use representations and models to communicate scientific phenomena and solve scientific problems. **Science Practice 3 (SP3)** The student can engage in scientific questioning to extend thinking or to guide investigations within the context of the AP course. **Science Practice 7 (SP7)** The student is able to connect and relate knowledge across various scales, concepts, and representations in and across domains.	**Essential knowledge 1.B.1 (EK1.B.1)** Organisms share many conserved core processes and features that evolved and are widely distributed among organisms today. **Essential knowledge 3.E.2 (EK3.E.2)** Animals have nervous systems that detect external and internal signals, transmit and integrate information, and produce responses.

Connections to the Nature of Science	
Vision and Change for Undergraduate Biology Education Core Competencies and Disciplinary Practices	A Framework for K-12 Science Education Understandings About the Nature of Science
Ability to Apply the Process of Science Understand the process of science and how scientists construct new knowledge by formulating hypotheses and then testing them against experimental and observational data.	**Scientific Knowledge is Based on Empirical Evidence** Science includes the process of coordinating patterns of evidence with current theory. Scientific arguments are strengthened by multiple lines of evidence supporting a single explanation.