SciTECH

Science | Technology | Engineering | Culture | Hacks

Vol. 1, Issue 2

HONEY BEES:
PRIME MODELS FOR MEMORY STORAGE AND ALZHEIMER'S P1

THE BEAUTIES OF PURE MATHEMATICS AND MUSIC

DOUBLE, DOUBLE, TOIL AND TROUBLE
The Future of Computing After Moore's Law • P6

Contents

Editorial Board

Cover by RND Systems.
Back Cover from Choate Rosemary Hall.

HONEY BEES

PRIME MODELS FOR MEMORY STORAGE AND ALZHEIMER'S

by KATE SPENCER '20

In an effort to further understand the long-term development of memory and memory impairment diseases, researchers at the University of Queensland in Australia recently used honeybees to uncover the impact of DNA methylation. By examining the presence and significance of DNA methylation, the study found that this particular epigenetic process does in fact affect the way a honeybee's brain remembers and relearns.

The study, conducted by Dr. Stephanie Biergans, depicted that certain molecular mechanisms that regulate memory specification and relearning could have a significant impact on the ways experiences are integrated into long-term memory. Biergans states, "There is thought to be a genetic predisposition for some conditions, such as Alzheimer's and dementia, but in many cases environmental factors determine whether the disease will manifest." The researchers aim to use honeybees in order to discover more about molecular and environmental changes that lead to memory impairment diseases.

Although honeybees may appear to be a questionable parallel to Alzheimer's patients, they are valid models for helping researchers discover more behind the disease's cause. "Honeybees have an amazing capacity to learn and remember," says Biergans. "They can count up to four, and orientate themselves by learning patterns and landmarks. They are also social insects that interact, teach, and learn, making them successful foragers. Bees remember how to find a food source, how good the source was, and how to return to the hive." Because they can form complex memories like humans but have a simple brain structure, honeybees have become a prime model for researching the formulation of long term memories.

Researchers in the past have established that Alzheimer's disease and memory formation are partly controlled by molecular changes in the brain's chemistry. As the DNA makeup is altered, physical changes may occur, including differing or new neutral activity and connections. Epigenetic mechanisms, which are a series of molecular changes that can occur due to experience or environmental changes, are known to affect memory formation in humans. They regulate gene expression through DNA modification without changing the individual gene. The Australian research team wanted to determine if DNA methylation, an epigenetic process, had an impact on how honeybees learn and relearn data. Based on what is needed, methylation can make a certain gene stay active or inactive. Biergans added, "We knew that DNA methylation is an epigenetic process that occurs in the brain and is related to memory formation. When we block this process in honeybees, it affects how they remember."

The experiment involved two separate groups of honeybees, both of which were taught to expect sugar in the presence of a particular scent. The first group learned over an extended period of time; however, the second

Bee brains could give scientists clues about how humans create memories—and how to stop dementia and Alzheimer's. *REUTERS.*

group was only exposed to the scent once. As explained by the research team, they taught the bees, "to associate an odor with a sugar reward, similar to the olfactory learning taking place when a bee collects nectar from a flower during foraging." Then they blocked DNA methylation in a few bees in each group using an inhibitor compound. After evaluating both groups with and without DNA methylation, the scientists changed the smell and repeated the process in order to find stable and ongoing conclusions.

The University of Queensland's researchers ultimately found that DNA methylation plays a significant role in defining how a bee can relearn. "When the bees were presented with sugar and a smell many times together, the presence of DNA methylation increased memory specificity – they were less responsive to a novel odor," summarizes Biergans. "On the other hand, when only introduced to the combination once, DNA methylation decreased specificity." The results reflect how a honeybee gets and chooses its source of nourishment; when a flower continuously proves to be a proficient source of food, the bees will search for that specific smell. Bees that were not able to methylate were not able to form memories that were as strong or specified as the bees that did methylate.

Through this process, the scientists have discovered information that may allow them to provide treatments for brain diseases in the future. The simple mind of a honeybee has proven to be an accurate and helpful model in understanding the processes behind long-term memory formation. Linking the physical changes in neural connections to human brains could uncover the mystery behind conditions like dementia and Alzheimer's. Future research is expected to build off of Biergans' study.

Researchers are hopeful to expand the scientists' findings. Biergans states, "By understanding how changes to the epi-genome accumulate, manifest, and influence brain function, we may, in the future, be able to develop treatments for brain diseases that also develop over a lifetime." •

BIBLIOGRAPHY

1. Glowatz E. Honeybee Brains Give Clue for Alzheimer's Cure. Medical Daily. 2016 Dec 9 [accessed 2017 Feb 13]. http://www.medicaldaily.com/alzheimers-cure-honeybee-brains-and-memories-give-clue-dementia-treatment-406124
2. Honeybee memories: Another piece of the Alzheimer's puzzle? ScienceDaily. [accessed 2017 Feb 17]. https://www.sciencedaily.com/releases/2016/12/161208125919.htm

TUMOR IS UNEARTHED FROM FOSSILIZED MAMMAL

A zoomed in section of the odontome, each resembling a miniature tooth. *Christian Sidor/Megan Whitney*

by SKYLAR HANSEN-RAJ '20

When two paleontologists working at the University of Washington cut into the fossilized jaw of a gorgonopsian, they got more than what they bargained for: a benign tumor composed of small toothlets. As they reported in a letter published on December 8th, 2016 in the Journal of the American Medical Association Oncology, the team discovered several tooth-like structures embedded in the fossil's jaw. This is evidence for compound odontoma, a common tumor in mammals today. But this creature existed almost 255 million years ago—before mammals were even around. "We think this is by far the oldest known instance of a compound odontoma," said senior author Christian Sidor, a professor of biology and curator of vertebrate paleontology at the Burke Museum of Natural History and Culture. Indeed, before this discovery, the oldest known case of odontoma was found in the fossils of mammals from the Ice Age. Sidor says that the discovery "would indicate that this is an ancient type of tumor."

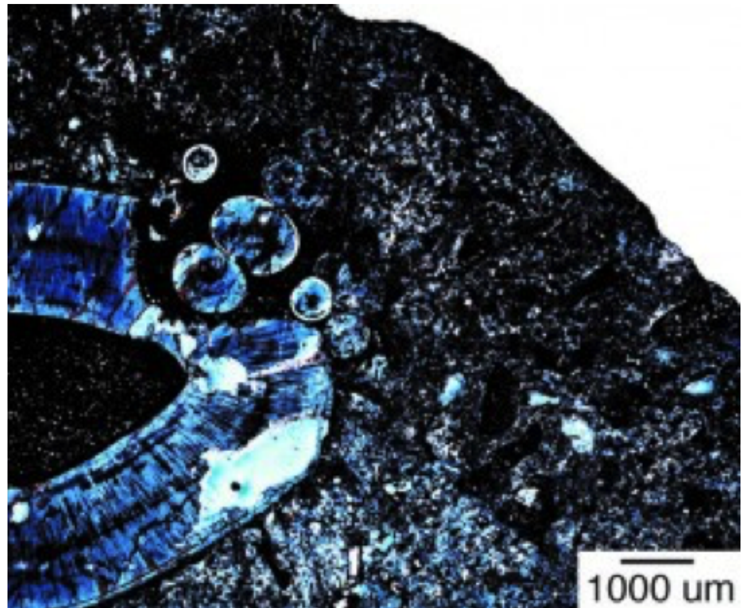
Compound odontoma still exists in mammals today. Odontomas are characterized by a mass of small "toothlets" amalgamated together along with dentin and enamel. They grow within the gum and other soft tissues of the jaw, causing pain and swelling. They can also disrupt the position of surrounding teeth. Because these toothlets do not spread to other parts of the body, they are considered benign. However, doctors today opt to remove them via surgery to stop the pain.

Surgery was not an option for the creature studied with this tumor. The fossil was once a gorgonopsian. Gorgonopsians were carnivorous, land-based vertebrates that lived around 270-250 million years ago. Their fossils have been found all the way from Russia to Africa. Gorgonopsians are part of a larger group of animals called synapsids. Synapsids are sometimes referred to as "mammal-like reptiles" because the now-extinct synapsids possess some features of mammals. This means they are more closely related to mammals, such as humans, than they are to reptiles. "Most synapsids are extinct, and we — that is, mammals — are their only living descendants," said Meghan Whitney, lead author of the experiment. "To understand when and how our mammalian features evolved, we have to study fossils of synapsids like the gorgonopsians." Whitney decided to study the teeth to see if they were similar to teeth in mammals today.

In order to examine gorgonopsian teeth, Whitney cut into the fossilized jaw. She states, "Most reptiles alive today fuse their teeth directly to the jawbone," said Whitney. "But mammals do not: We use tough, but flexible, string-like tissues to hold teeth in our sockets. And I wanted to know if the same was true for gorgonopsians." She, along with Larry Mose, an undergraduate student working with her, decided to cut into the jaw. Although a controversial move, there was no other way to prove the similarities or differences through a purely physical examination. Since this would damage the fossil, the pair used a solitary gorgonopsian lower jaw that Sidor had collected in southern Tanzania.

After cutting through the fossil for closer examination, Whitney and fellow researchers found a benign odontoma. Once Mose prepared the thin slices of the gorgonopsian jaw, he and Whitney immediately noticed something unusual. Irregular clusters of up to eight tiny, round objects were embedded next to the root of the canine. Under the higher magnification of the microscope, the paleontologists could distinguish the small objects as toothlets, which even harbored distinct layers of enamel and dentin. “At first we didn’t know what to make of it,” said Whitney. “But after some investigation we realized this gorgonopsian had what looks like a textbook compound odontoma.” Since the fossil has been dated at 255 million years old, this makes it the oldest case of odontoma ever reported, and the first case in a non-mammal. The discovery of an odontoma in a gorgonopsian indicates that this “mammalian condition” has existed well before the first mammals had evolved. “Until now, the earliest known occurrence of this tumor was about one million years ago, in fossil mammals,” said Judy Skog, program director in the National Science Foundation‘s Division of Earth Sciences, which funded the research. “These researchers have found an example in the ancestors of mammals that lived 255 million years ago. The discovery suggests that the suspected cause of an odontoma isn’t tied solely to traits in modern species, as had been thought.”

This research proves the oldest occurrence of odontoma in a mammal relative, implying that the odontoma hasn’t changed much in the past 255 million years. It also shows how disease can be studied from the past, and helps guide future scientists to study the beginnings of tumors and where they may have come from. Paleontology can contribute to medicine by shedding light on the history of disease. Sider notes, “This discovery demonstrates how the fossil record can tell us a lot about our present-day lives—even the diseases or pathologies that are part of our mammalian heritage.” •



Odontoma found in the gorgonopsian’s lower jaw
The dark area (right) is bone. The backwards “C” shape is the canal root. The small circular objects are “toothlets.” *Megan Whitney/Christian Sidor/University of Washington.*

Partial gorgonsopic lower jaw.
This not the specimen in which the odontoma was discovered. *Christian Sidor/University of Washington.*

BIBLIOGRAPHY

1. Sidor C, Whitney M. Frequently asked questions: odontoma in a gorgonopsian. UW Today. [accessed 2017 Feb 13]. http://www.washington.edu/news/2016/12/08/frequently-asked-questions-odontoma-in-a-gorgonopsian/
2. Urton J. Fossilized evidence of a tumor in a 255-million-year-old mammal forerunner. UW Today. [accessed 2017 Feb 13]. http://www.washington.edu/news/2016/12/08/fossilized-evidence-of-a-tumor-in-a-255-million-year-old-mammal-forerunner/
3. Whitney MR. Odontoma in a 255-Million-Year-Old Mammalian Forebear. Oncology | JAMA Oncology | The JAMA Network. 2016 Dec 8 [accessed 2017 Feb 13]. http://jamanetwork.com/journals/jamaoncology/article-abstract/2591155
4. Fossilized evidence of a tumor in a 255-million-year-old mammal forerunner. ScienceDaily. [accessed 2017 Feb 13]. https://www.sciencedaily.com/releases/2016/12/161208143408.htm

The Genetic Link to Personality Traits and Mental Disorders

by LILY DING '20

Scientists have already started to link distinct personality traits and psychiatric disorders to particular regions of the human genome. When asked about how this step forward in the genome-wide association studies (GWAS) community will affect his work, Dr. Jeff McCrae, a personality psychologist at the National Institute on Aging (NIA), says,"[Genetic markers] could provide one more objective indicator against which to evaluate our instruments." While discovering actual "personality" genes is decades away, identifying regions of the genome that are correlated with certain personalities is a huge step in personality research.

A new meta-analysis of GWAS has established six loci, or regions, that are connected with several different personality traits and disorders such as introversion, neuroticism, and agreeableness. Researchers at the University of California San Diego School of Medicine have found the connection between different parts of the genome and personalities.

The goal of this research is to discover which genes affect brain function that, in turn, affect the way different people interact with their environment. The complex way that these genes interact with one another influences various traits all at once in subtle ways. This is why progress is slow in behavioral genetics research. Any one gene likely accounts for only a one or two percent variance in a specific personality trait. Senior author Chi-Hua Chen, PhD, assistant professor in the Department of Radiology at UC San Diego School of Medicine says, "Although personality traits are heritable, it has been difficult to characterize genetic variants associated with personality until recent, large-scale GWAS."

Even so, more and more researchers are diving into the unknown and complex fray of behavioral genetics, hopeful that they will not only be able to identify how genes relate to personality, but also gain an understanding about what makes people tick. Undeterred by this seemingly impossible task, the researchers at the University of California San Diego School of Medicine have developed a comprehensive understanding of personality and psychopathology, in addition to the fragile balance between genes and their effect on one's personality.

There are five psychological factors commonly used to measure differences in personality. The first is extraversion versus introversion, which reflects how assertive, talkative, and integrated an individual is. The next is neuroticism versus emotional stability, which measures anxiety and depression along with other traits of deteriorating mental health. Furthermore, agreeableness and antagonism measure the cooperativeness and compassion of an individual. Lastly, how open one is to new experiences suggests intellectual curiosity and creativity. In the past, meta-analyses have attributed around forty percent of variance in personalities to genetic factors. GWAS, on the other hand, have discovered several variants associated with the five psychological factors.

In their new paper, Chen and her colleagues analyzed genetic variations among five different personality traits and six psychiatric disorders using data from 23andMe, a private genomics and biotechnology company, and other genetic companies.

Their results revealed that personality traits may be genetically similar to psychiatric disorders. In order to test this, they compared every psychological factor between personality traits and psychiatric disorders. Chen notes, "We found genetic correlations between personality traits and psychiatric disorders, but specific variants underlying the correlations are unknown." Researchers found that extraversion was associated with variants in the WSCD2 gene and near the PCDH15 gene; neuroticism was associated with variants on chromosome 8p23.1 and the L3MBTL2 gene. There were many genetic correlations between extraversion and attention deficit hyperactivity disorder, also known as ADHD. There was also a strong connection between openness and mental disorders like schizophrenia and bipolar disorder. Neuroticism, on the other hand, was genetically linked with depression, anxiety, and other internalized psychopathologies.

Looking to the future, social psychologist Dr. Ed Diener says, "Understanding the genes and their interactions will most certainly also help us understand environmental influences. We will be able to see when the environment 'overrides' the genes and why. And we will be able to see how environmental variations interact with genetic variations." •

Double, Double, Toil and Trouble:

The Future of Computing After Moore's Law

by HALEY CHANG '18

Responsible for successfully predicting the rate of technological revolution in the Information Age, Moore's law is now slowly coming to an end. The widely-credited exponential curve has governed the semiconductor industry since the 1960s, when Gordon Moore, who would later become one of the founders of Intel Corporation, published a paper in 1965 predicting a doubling every year in the number of components per integrated circuit — and that this trend would continue for at least another decade. Moore revised the forecast in 1975, postulating that the rate of doubling will reduce from once a year to once every two years.

In accordance with the revised Law, the first crude home computers of the 1970s transformed into the more sophisticated machines of the 1980s and 1990s, the computational speed became remarkably faster in the 2000s, and the 12.3-millimeter (mm) iPhone 3GS in 2009 slimmed down to the 7.10-mm iPhone 7 in 2017. These accomplishments were made possible largely because transistors have the quality of getting better as they get smaller. In other words, by decreasing the size of a transistor, it can be turned on and off with less power and at greater speeds than a larger one.

But many speculate that this fast-paced progress will soon slow to a halt. The talk about the demise of Moore's law is not a complete surprise, however; ever since its announcement, people have always been predicting the end of Moore's law, and engineers have always been coming up with novel ways to keep the pace of progress alive. In fact, Intel, which has been the leading maker of microprocessors for decades, helped turn this "law" into a self-fulfilling prediction. Since the 1990s, the industry introduced a strategy sometimes called More Moore, in which they release a research road map every two years to coordinate what its hundreds of manufacturers and suppliers should do to preserve the law.

Until a few years ago, Intel was able to consistently reduce the scale of the architecture of chip designs — but the cycle of "two years" has been lengthening. According to a 2017 Fortune article, it took Intel 23 months to get their chips from 65 nanometers (nm) to 45 nm, 27 months to get from 45 nm to 32 nm, 28 months to get from 32 nm to 22 nm, and another 30 months to shrink to the 14 nm chip released in September 2014. With more than two years already passed since then, Intel has yet to release its widely-anticipated 10 nm chip, which they hope will dispute some of the concerns of the industry that Moore's Law is slowing down.

In the past, the ease of doubling the computing power engendered a lack of incentive to experiment with other kinds of improvement. However, William Holt, who leads Intel's technology and manufacturing group, said last year that for chips to keep continuing their exponential trend of improvement, Intel will soon have to start using fundamentally new designs and materials to make transistors amenable to more shrinkage.

Below are some of the candidates for this inevitable change:

Graphene: Graphene, single-atom sheets of carbon, has electronically useful properties that remain stable under a wide range of temperatures like silicon. However, researchers face a paramount challenge in making graphene transistors a reality; graphene does not have a bandgap, a quantum property that makes it possible to turn a transistor from on to off. The next step would be to develop a method to open a sizeable and well-defined bandgap in graphene.

Carbon Nanotubes: Unlike graphene, carbon nanotubes develop a band gap and, along with it, several other semiconducting properties. However, carbon nanotubes are delicate structures — if a nanotube's diameter or the angle at which the carbon atoms are "rolled" varies by even a small amount, its band gap may vanish, rendering it useless as a digital circuit element. There is also the practical concern of being able to place nanotubes — billions of them — into neat rows with just a few nanometers apart.

Quantum Mechanics: One idea is to employ quantum mechanics to perform calculations much faster than any classical computer could ever do. By exploiting the quantum effects of superposition and entanglement, electronics harnessing the power of quantum mechanics may be quadratically or polynomially faster. D-Wave, a popular quantum annealer, and universal gate quantum computing are both frequently in quantum computing. •

The "P-versus-NP" Problem:

Where Are We Going, Where Have We Been?

Part 1 of 2: A Beginner's Guide

by JESSICA SHI '17

The computer science nerds of the world may be able to derive a fair amount of enjoyment from a website called "The P-versus-NP Page," a mind-numbingly long list of papers — whose publication dates range from the late 1980s to just a few months ago — detailing attempted advancements and solutions to the titular problem. Of these 116 entries, exactly one has been broadly verified and accepted by experts in the research community: it shows, as the website describes, "that a certain approach to settling this question will never work out."

In fact, the P vs. NP problem is so difficult that the Clay Mathematics Institute named it one of seven "Millennium Prize Problems," each of which has a one million dollar monetary award attached for the first submission of a complete and correct solution. But before readers who have yet to hear of this problem walk away, let us first try and understand the dilemma at hand: what is "P," what is "NP," and why does anyone want to compare them?

Although it was not posed in its current form until 1971, the P-versus-NP problem has origins in letters from John Nash to the National Security Agency and from Kurt Gödel to John von Neumann two decades prior. There are many equivalent ways to consider the question, but to put it simply: there are P problems, which can be easily solved by computers, and NP problems, which cannot be easily solved but can be easily verified once a solution is presented. (If you have a hard time conceptualizing how this can be, consider this analogy, published in *MIT Technology Review*. "Imagine a jigsaw puzzle: finding the right arrangement of pieces is difficult, but you can tell when the puzzle is finished correctly just by looking at it.")

All P problems are NP problems, but the question lies in whether these problem types are equal. That is, all easily solvable problems are definitely easily verifiable, but are all easily verifiable problems also easily solvable? As of now, mathematicians and computer scientists believe the answer is probably no (which would mean that P does not equal NP), but no one has been able to conclusively prove this to be true — or false.

Of course, a natural question may arise: what is defined as "easily" solvable? P actually stands for "polynomial time," and conversely, NP stands for "nondeterministic polynomial time." The notion of polynomial time helps us define which problems are easy and which are not — P problems can be solved with an algorithm that has execution time proportional to N^k for some non-negative integer k, where N describes the input size. NP problems cannot. The difference? Non-polynomial time algorithms can take a really, really, really long time. In fact, a lot appear to require brute-force methods — try every solution until you find one that works. (As a comparison, an algorithm with execution time proportional to N^3 with N = 100 will take three hours; an algorithm with execution time proportional to 2^N with N = 100 will take 300 quintillion years.)

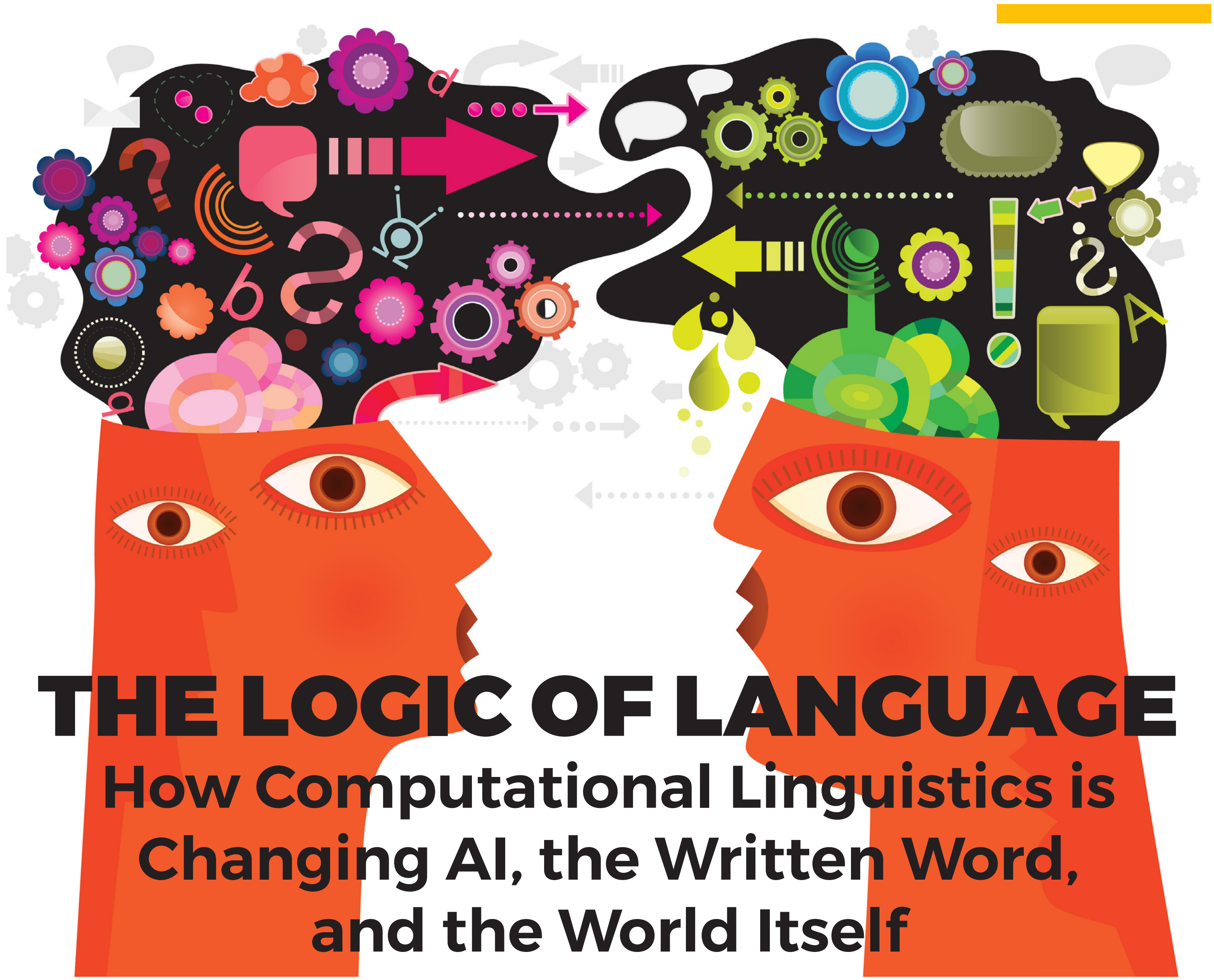
If, as many believe, P does not equal NP and this were proved to be the case, the revelation might not change much, since most things already operate under this assumption. However, if it were somehow to be shown that P *equals* NP, it would mean that many problems that are seemingly very difficult must also have some shortcut. This result would make the current progress in fields like cryptography invalid in some respects, but to greater magnitudes, it would also open up enormous possibilities: transportation could be scheduled more time- and cost-efficiently; manufacturers could improve their production rates and quality; learning-based tasks like visual recognition and language comprehension could be solved with little to no effort.

But of course, as MIT complexity researcher Scott Aaronson warns, don't get your hopes up. "There are good reasons why very few people believe that P equals NP," he says. "If it did, we'd be living in a fundamentally different universe, and we'd probably have noticed by now." •

Next term, we will go in-depth into specific methods used to attempt P vs. NP proofs and the implications of what has been learned so far.

BIBLIOGRAPHY

1. *Fortnow, L. The Status of the P versus NP Problem. Communications of the ACM [Internet]. 2009 [cited 2017 Feb 15]. Available from: http://cacm.acm.org/magazines/2009/9/38904-the-status-of-the-p-versus-np-problem/fulltext .*
2.
3. *Hardesty, L. MIT News [Internet]. Cambridge (MA): MIT; 2009 Oct 29. Explained: P vs. NP; [cited 2017 Feb 15]. Available from: http://news.mit.edu/2009/explainer-pnp.*
4.
5. *Miessler, D. Daniel Miessler Blog [Internet]. San Francisco (CA): Daniel Miessler; 2017 Feb 7. P vs. NP Explained; [cited 2017 Feb 15]. Available from: https://danielmiessler.com/study/pvsnp/#gs.dyqnx3Y .*
6.
7. *Pavlus, J. MIT Technology Review [Internet]. Cambridge (MA): MIT; 2010 Aug 19. What Does 'P vs. NP' Mean for the Rest of Us?; [cited 2017 Feb 15]. Available from: https://www.technologyreview.com/s/420290/what-does-p-vs-np-mean-for-the-rest-of-us/.*



THE LOGIC OF LANGUAGE

How Computational Linguistics is Changing AI, the Written Word, and the World Itself

by ZEV NICOLAI-SCANIO '18

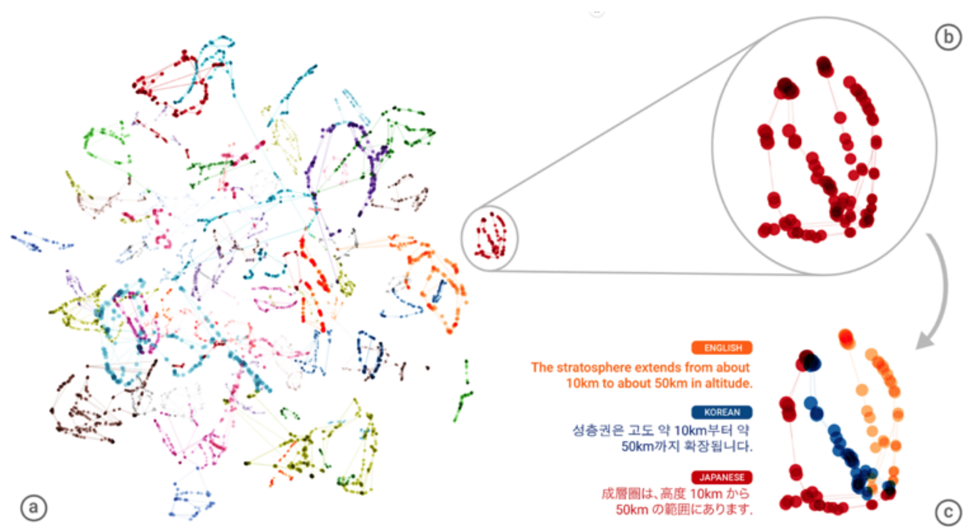
Written language has been central to computer science since the beginning of the discipline, even before computers, at least in their modern form, physically existed. In 1950, Alan Turing — the legendary inventor of Bletchley Park's "Bombe" codebreaker and the philosopher behind the "Turing machine," a central idea of automata theory — formulated one of the most important challenges of the nascent field of artificial intelligence (AI). It would come to be known as the "Turing test." The premise of the challenge was simple: could a computer running an AI algorithm engage in a text-based conversation with a human being and be indistinguishable from another human participant over the course of the conversation?

The pursuit of this challenge expanded to become its own subfield of AI, Natural Language Processing (NLP). As human-computer interactions became more and more common in everyday life, the pressure for advancements in the field of Natural Language Processing rose. For example, search engines needed to process the text queries of their users; automated translation services needed semantic, structural, and grammatical understanding of language; and software companies entered the race to produce ever more intricate virtual assistants for their users. What had started as an academic challenge became the center of a multibillion dollar industry.

The industry's AI experts were ready to tackle the problem. The algorithmic toolkit of AI had advanced tremendously since the days of Turing, and new techniques of machine learning had revolutionized image processing. The most famous of these budding algorithmic families, artificial neural networks, were hyped as a panacea for AI woes. Yet advancements in NLP remained elusive.

The sticking point was the data representation of language. The building blocks of images, namely pixels, are ideally suited for processing with artificial neural networks, as well as the techniques of machine learning in general. However, the building blocks of language, namely letters, are not as accommodating.

Think of it this way. If I were to chop a photograph into sets of pixels, it would be easy to identify the features on these sets — edges, lines, colors, squiggles — and then, by making a map of these features and their locations, one can produce an outline of the image. Now imagine if I broke The Great Gatsby into sets of letters and gave those to you. What could you "sketch" about the book? Not much. Language is abstract and sequential. In contrast to pixels, letters are not a data format but rather a human construct, and thus the "information" represented by different letters in different scenarios varies based on the etymology of the given language. In addition, because language is sequential, information stored in a given set of letters is dependent on the set before and, sometimes, the set after the target set[1].



A 3-dimensional representation of internal network data at Google
(a) A visualization of the word vectors of a few sentences. (b) A close-up view of just one sentence. (c) The word vectors for the same sentence in English, Korean, and Japanese all plotted simultaneously. The fact that similar words are grouped spatially across all three languages is a demonstration of the properties of word vectors that allow for their use in translation. *Google.*

Methods of computational linguistics offer a new way to approach the problem of the data representation of language. The most promising is a technique known as "word vectorization" or "word embeddings." The goal of this technique is to produce a novel data representation of language called "word vectors." Word vectors on a fundamental level are just like any vector in that they are composed of a set of components with one component value for each axis (also known as a basis vector). The trick behind word vectors is the vector space used. Instead of the familiar (x, y, z) coordinate system, word vectors reside in spaces of a hundred, or often a thousand, dimensions. This higher dimensional space acts to store the word's meaning in a manner very similar to that of a human reader. Each of the hundreds upon hundreds of basis vectors can be thought of as representing a certain concept. These can be as simple as "singular" and "plural" or as abstract as "idyllic" and "royal."[2]

Vectors not only provide a machine friendly way to represent language, which has applications across the field of NLP, but also hold two other interesting properties as well. The first is that they allow for a sort of "word algebra." This means that the vector sums and differences of different words reflect their relationships in language: "cats" minus "cat" is equal to "dogs" minus "dog" is equal to the plural of a noun minus the singular of a noun in general. Computational linguists soon discovered other patterns similar to these. For example, computational results of a capital city minus the name of its country were consistently equal. One famous word vector equation that made headlines in the computational linguistics community was that "king" minus "man" plus "woman"

1. These problems exist in character based languages where words are the building blocks as well, just to a lesser extent.

2. All that is algorithmically required is that the basis vectors are mutually orthogonal. Therefore, each of the basis vectors does not necessarily correspond to what a human would consider a concept. What is important is that all this higher level conceptual information of language is successfully represented in the vector.

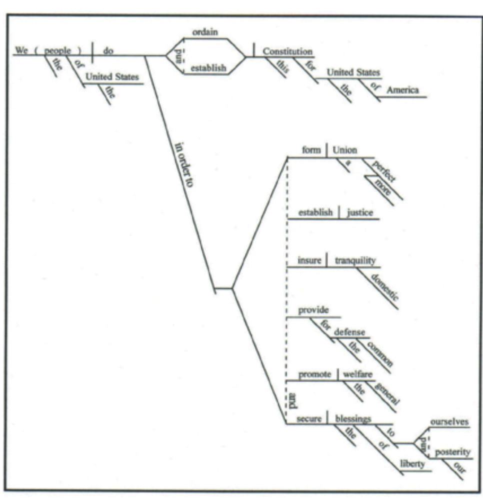
equaled “queen”. Another interesting property of word vectorization is that since the word vectors encode the meaning of the words at an abstract level, they are independent of the language of the given words. This means that by looking at the word vectors of words in one language and finding the words from another language with the most similar word vectors, one can create an efficient method for translation. In fact, Google recently updated Google Translate to use a new word vector based algorithm.

In addition to the elegant logical and mathematical appeal of word vectors, using them at scale is reasonable. Efficient algorithms exist to quickly compute the vector representations for large numbers of words in a language using only a reasonably sized database of text sources. While the two main algorithms (CBOW and Skip-gram) differ in the details of the implementation, the central methodology is the same: use an artificial neural network to learn the vector representation of a given word based on the other words surrounding it and for all of its occurrences in the text database.

Another place where computational linguistics is making a difference in NLP is an area known as “sentence compression”. The goal of sentence compression is to extract the pertinent information from a block of text. The interest in sentence compression stems from the necessity of this skill in allowing computers to answer questions in a manner similar to humans. Google is one of the leaders in the research of sentence compression techniques. For readers who are frequent Googlers, you might have noticed that the search engine now directly answers many questions in addition to returning search results. For example, “What is the flying speed of an unladen swallow?” returns “31 – 40 mph” and the question “Who is Batman?” returns “Batman is a superhero co-created by artist Bob Kane and writer Bill Finger and published by DC Comics…” This new feature is made possible by deep neural networks, one of the most powerful algorithms of modern AI. Yet deep neural networks, like all forms of machine learning, are only as good as their training procedure.

David Orr, a Product Manager at Google, has a team of computational linguists who built an optimal training procedure. First, they decided that the training would include two types of data, referred to as “silver” and “gold”. The “silver” category consists of data that is readily available. Orr and his colleagues decided to use news articles and their corresponding headlines as the “silver” set of training data. The reasoning was that headlines concisely summarize news stories and therefore would provide the deep neural network with training examples of sentence compression. The “gold” set of training data was custom-made by the Google’s Pygmalion team, a group of computational linguists employed expressly for this purpose. The Pygmalion team selected choice snippets of text and then manually entered what the sentence compression result should be, thus providing the input output pairs needed for training the deep neural network. But what made this “gold” data special was that the Pygmalion team also included the grammatical function of each word in the text. This added layer of information improved the learning process and results.

Just as Ovid’s character, Pygmalion, who through the power of the gods brought to life the elegant ivory form of his greatest statue, computational linguists around the world, through the power of computer science, are bringing to life the inner logical form of human language. •

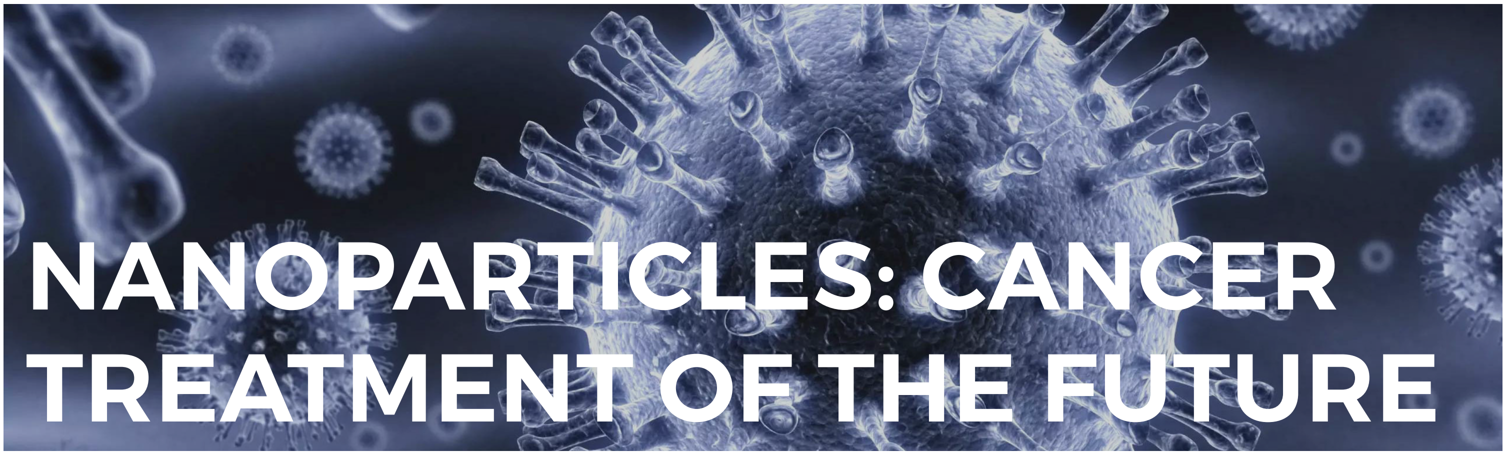


This a diagram representing the type of grammatical and structural information included for every sentenced used by the Pygmalion team. *Eugene R. Moutoux of German-Latin-English.*

BIBLIOGRAPHY

1. Colah. Deep Learning, NLP, and Representations. colah’s blog. July 7, 2014 (accessed Feb 8, 2017) from: http://colah.github.io/posts/2014-07-NLP-RNNs-Representations/.
2. Colyer A. The Amazing Power of Word Vectors. the morning paper. April 21, 2016 (accessed Feb 8, 2017) from: https://blog.acolyer.org/2016/04/21/the-amazing-power-of-word-vectors/.
3. Grimes S. An Inside Update on Natural Language Processing. Breakthrough Analysis. June 23, 2016 (accessed Feb 8, 2017) from:https://breakthroughanalysis.com/2016/06/23/jbnlp/.
4. Metz C. Google’s Hand-Fed AI Now Gives Answers, Not Just Search Results. Wired. Nov 29, 2016 (accessed Feb 8, 2017) from: https://www.wired.com/2016/11/googles-search-engine-can-now-answer-questions-human-help/.
5. Schuster M, Johnson M, Thorat N. Zero-Shot Translation with Google’s Multilingual Neural Machine Translation System. Google Research Blog. Nov 22, 2016 (accessed Feb 8, 2017) from: https://research.googleblog.com/2016/11/zero-shot-translation-with-googles.html.
6. Wikipedia. Turing test. https://en.wikipedia.org/wiki/Turing_test (accessed Feb 8, 2017).

Cover illustration by iStockPhoto.



NANOPARTICLES: CANCER TREATMENT OF THE FUTURE

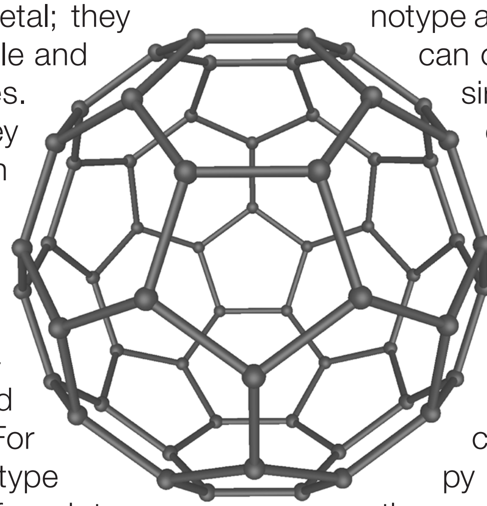
by ALYSSA SHIN '18

Most people, when they think of cancer, shudder at the thought of the harrowing chemotherapy and radiation therapy that many cancer patients have to endure.

Chemotherapy and radiation therapy are notoriously taxing on the body because they are non-specific on the cellular level. This means that they kill cancer cells but also normal, healthy cells. Cancer cells divide rapidly and often metastasize, spreading from a localized area throughout the body. Chemotherapy uses this characteristic of cancer cells and targets rapidly dividing cells. This method of treatment also destroys healthy cells that replicate quickly, such as hair cells or liver cells. As a result, patients often experience side affects such as diarrhea, hair loss, or nausea.

Radiation therapy, on the other hand, targets localized tumors. However, it also kills normal cells in the surrounding area. Thus, while chemotherapy and radiation therapy both wipe out cancer cells and tumors, they often greatly harm the patient in the process. So, how can we target cancer cells specifically while leaving healthy cells intact?

Nanoparticles may provide the answer to this question. Nanoparticles can be made of organic molecules or even metal; they are also highly versatile and manipulable structures.



First, because they are so small, they can be injected into the bloodstream with minimal invasiveness. Second, their shapes and surfaces can be engineered to find cancer cells. For example, if a specific type of cancer cell was found to overexpress a certain receptor, the nanoparticle created to target those cancer cells would be created with the corresponding ligand, or the partner binding structure, to allow it to bind to the receptors on the cancer cell surface and thus find the cancer cells.

Therefore, unlike chemotherapy and radiation therapy, nanoparticles can identify and selectively eliminate cancer cells. Third, nanoparticles can carry altered genes, which act as delivery vehicles, depositing them at the target cell or target site. This means that scientists can change the genetic makeup of the cell, allowing scientists to produce virtually any phenotype as long as they can create the desired sequence of genes.

While nanoparticles could resolve all of the side effects associated with chemotherapy and radiation therapy, we must be careful in utilizing this technology. Before nanoparticles are implemented in clinical use, extensive testing must be carried out to ensure the safety of the patients.

In addition, when using nanoparticles, specifically for gene therapy, we must make sure that the nanoparticles do truly target cancer cells specifically in order to prevent wrong delivery, which could lead to disastrous mutations. •

Cover image from the National Cancer Institute.

A Peak into Drone Navigation and Obstacle Avoidance

by WILLIAM WU '20

Drones, or unmanned aerial vehicles, have experienced a high degree of popularity in recent years for research, aerial filming, and racing Many researches have improved quadcopters over the years so that drones can be piloted more easily. Major goals in drone research include navigation and obstacle avoidance. The article aims to demystify the existing technologies behind navigation and obstacle avoidance of drones on the market, and provide some knowledge in order to better understand the current state of drone research.

Current research has focused on quadcopters, or "quads". Although they can be built inexpensively, quads are not only effective machines but are also extremely versatile, lending themselves to several different fields, such as aerial filming. With only four motors, quads can effectively roll, pitch, and yaw. The most notable producer in the civilian quadcopter market is DJI. Founded in 2010, DJI uses rigorous testing and research in order to create the most efficient drones.

Since the first generation of Phantom, DJI's flagship drone brand, the quads have utilized automatic GPS. Automatic navigation takes over under when the drone is either out of range from the controller or when the "return to base" command is given. Under both circumstances, the drone will mark the last recorded location of controller as its destination, navigate to that location, and land under the guidance of GPS. Though this method is the most prevalent and longstanding, it comes with a number of disadvantages. The primary drawback is that outdoor environments come with obstacles that are difficult to predict, and these variables in the environment, like trees and telephone poles, may cause the drone to crash. DJI has taken this into their account when designing new drones. Starting with Phantom 4 series, DJI released a powerful new feature: obstacle avoidance. This function dramatically decreases the possibility of drone crashes through using sensors on board, and the drone uses these to sense its surroundings and warn the user of obstacles. The drone may also come to a stop if an obstacle is too close, mitigating the possibility of collision.

The algorithm that DJI uses is called optical flow, which aids the drone in being able to map the surroundings during cameras or flow sensors There are many existing algorithms for optical flow, and McGuire et al. presented one of them, Edge-FS, as a robust velocity estimation algorithm. The algorithm was proven to be efficient enough so that it can be run on a pocket-sized drone with limited computing power. Essentially, the algorithm analyzes the image from two points: $t = k$ and $t = k - n$, where $t = k$ represents displacement at time 0 and $t = k – n$ accounts for motion after the beginning. The algorithm first analyzes the images, and then matches the images with certain visual cues. With the calculated physical displacement and the measured time displacement (n), velocity can be derived. This algorithm applies to obstacle detection as well. With the aid of optical flow, a basic control model with four states can accomplish automatic obstacle avoidance. The quadcopter begins with "check mode", in order to determine whether or not there are obstacles in the way, and the quadcopter will then switch to "move mode", in which the drone moves with a certain velocity. When obstructions are detected within a certain range, the quadcopter will instantly come to a stop and turn a certain number of degrees to avoid the obstacle. If the obstruction is cleared, then the quadcopter will resume moving forward.

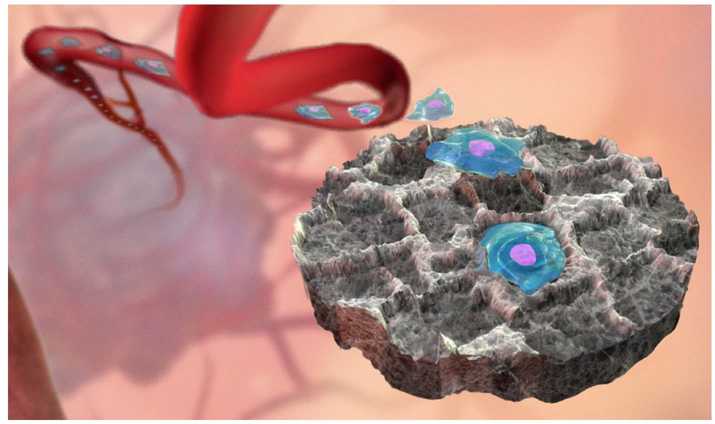
The optical flow system can also be applied to target tracking, obstacle detection, and avoidance. The Phantom 4 series fully utilizes optical flow and, using two arrays of sensors on board separate from the camera, target tracking and obstacle avoidance can all be accomplished. Additionally, users can now command the drone to fly to a certain place automatically, or track a certain object with optical flow.

In conclusion, this article presents two methods of drone navigation and obstacle avoidance: GPS navigation, and optical flow system. GPS navigation is an old way to navigate, and it presents problems such as crashing because of its inflexibility. The appearance of optical flow efficiently solves the problems of navigation, and it allows for more possibilities, including object tracking and detection. Through future research, the accuracy of both GPS navigation and optical flow can be improved, making it safer and easier for drones to fly. •

NANOTECH SCAFFOLDS

BUILDING THE BODY'S DEFENSES AGAINST CANCER

by IMAD RIZVI '18



Conceptual illustration of implantable sponge-like device attracts cancer cells from the bloodstream.
University of Michigan.

Although cancer is still the second leading cause of death in the US, cancer-related deaths have been steadily decreasing over the past few decades with the help of preventative care options that identify cancers early when they are easier to remove and therefore cure. A recent study from the University of Michigan shows how an engineered scaffold inserted under the skin can offer a new preventative option to detect cancers and slow their development.

The small device takes advantage of interactions between an organism's immune system and the cancer. Cancer can be attracted to various beacons of immune cells once the cancer has taken over certain parts of the body. When the engineered device is inserted under the skin, immune cells naturally gather in the area and around the scaffold, as they do not recognize the implant as part of the body. Cancer cells are then attracted to the area. Medical imaging reveals when the cancer cells have localized in the scaffold, and the device can then be removed to study the colonies.

Scientists can examine these cancer cells to create a response that is precise and specific to the patient. Not only does this device allow scientists to study the cancer cells, but it also slows the spread of cancer around the body. In a mouse model, researchers at the University of Michigan found that after 5 days, cancer cells had been attracted to the device, yet none had spread to the lung, liver or brain. After 15 days, the study showed that there were 64% fewer cancer cells in the liver and 75% fewer in the brains of the mice tested. The results showed the potential for the engineered scaffold to slow the progression of cancer and limit its spread to critical areas of the body, which is where it can wreak the most havoc.

This new innovation offers many potential uses that can help the future of cancer treatment. For many patients, cancer is discovered late on when it has already spread to different areas of the body, making it difficult to treat. Preventative screening — being able to detect cancer before it spreads throughout the body — is a priority for doctors. Since the device has minimal negative impacts and attracts any cancer cells that may be in the body, it offers a solution for finding cancers early on in their development. The device could be used for patients with a family history that predisposes them to certain forms of cancer or for monitoring patients who no longer suffer from cancer and making sure it does not return. For breast cancer specifically, which is difficult to detect until it has spread to other parts of the body, the device offers a solution to a critical condition affecting millions around the world.

This engineered implant may not offer a cure for any types of cancer, but it is still a major innovation because of the potential it has in detecting and slowing the spread of the disease. As the device is made from FDA-approved material, the technology is still not ready for human use and is currently undergoing more tests to ensure its safety. When it is approved, the device could have the potential to save thousands of lives by catching cancer before it becomes too difficult to cure. •

Trump Rocks Relationship Between Government and STEM: A Timeline

by ELYSE CORNWALL '18

Although cancer is still the seconThough STEM did not constitute a major part of President Donald Trump's platform during the election, he has initiated profound and widespread change in the way the federal government approaches STEM topics. With his "America First" focus, President Trump aims to overhaul the components of Obama's environmental legacy that his administration deems harmful and unnecessary through reorganizing various national STEM organizations such as the Environmental Protection Agency (EPA), the National Aeronautics and Space Administration (NASA), and the Department of Energy (DOE).

The extent to which climate change solutions and research will be inhibited depends on how effectively Trump can unravel his precursors' environmental policies. For both those hopeful for and unsettled by the changes that Trump and his STEM policies will bring, this piece offers a detailed timeline of the various actions President Trump has taken regarding STEM thus far.

President Trump appoints Rick Perry as Secretary of Energy. A proponent of local fossil fuel extraction, the former governor of Texas holds beliefs consistent with Trump's mission to take advantage of all resources on (or under) American soil. As the possible head of DOE, Perry's mission contrasts heavily with those of the climate change and clean energy programs that the DOE enforced during the Bush and Obama administration. In fact, Perry had previously proposed that the Energy Department, while under the Obama Administration, be disassembled.

On the first day of Trump's presidency, the White House website announced that Obama's Climate Action Plan would be dismantled, shifting the DOE's focus away from alternative energy and climate change research. A statement issued on the White House Website contends that this new energy policy will create jobs, raise salaries, and boost the national economy. This announcement is consistent with Trump's view that global warming was "created by and for the Chinese in order to make U.S. manufacturing non-competitive." By removing the energy and emission regulations that were previously characteristic of the DOE, the Trump Administration will have access to national shale, oil, and natural gas resources that Obama's Climate Action Plan protected.

Over 800 scientists and researchers express their concerns about the environmental impact of the "America First" plan in an open letter addressed to President Trump. The writers outline six important steps pertaining to climate change that might aid in, "this nation's efforts to tackle this crisis." While an online petition in support of the letter has secured over 128,000 signatures, the scientific argument that the letter proposes has received no response from Trump's administration.

The Trump administration proposes to concentrate more of NASA's budget into deep space exploration, and consequently, decrease funding for the Earth Science Department. Satellites for NASA's Earth science research monitor global temperatures, clouds, and ice. President Trump's senior adviser on space policy, Bob Walker, described the publication of these reports as "politically correct environmental monitoring," as they produced information that supported global warming. Dr. Lawrence Krauss, theoretical physicist, expressed his fears of President Trump's plans in an article called "Donald Trump's War on Science," stating, "Science is the one domain in human life where bias and prejudice are systematically eliminated; now those very forces are set to undermine the practice of science in America." Krauss argues that the Trump Administration's motivation to defund this research is an attempt to push climate change awareness to the backburner. Beyond losing access to climate trends, Krauss adds, NASA would be unable to track extreme weather trends from space. Despite dissent from Krauss and other experts, Trump's STEM policies currently remain fixed. •

The Beauties of Pure Mathematics and Music

by SAM MARKOWITZ '17

What makes something sound beautiful? We can listen to two notes played simultaneously (called an interval) on the piano, and while many of these intervals sound pleasant to the ear, others may sound cacophonous. These are the notes that make you cringe, or those that sound like something out of John Cage's music. So what are the differences among the notes that make some intervals mellifluous and some not?

We know from physics that each of these notes has a certain frequency, and for simplicity, we will consider only the notes on one octave of the piano. In an octave, we find only 12 of the infinite number of possible frequencies. We can analyze these intervals to see how notes relate to one another.

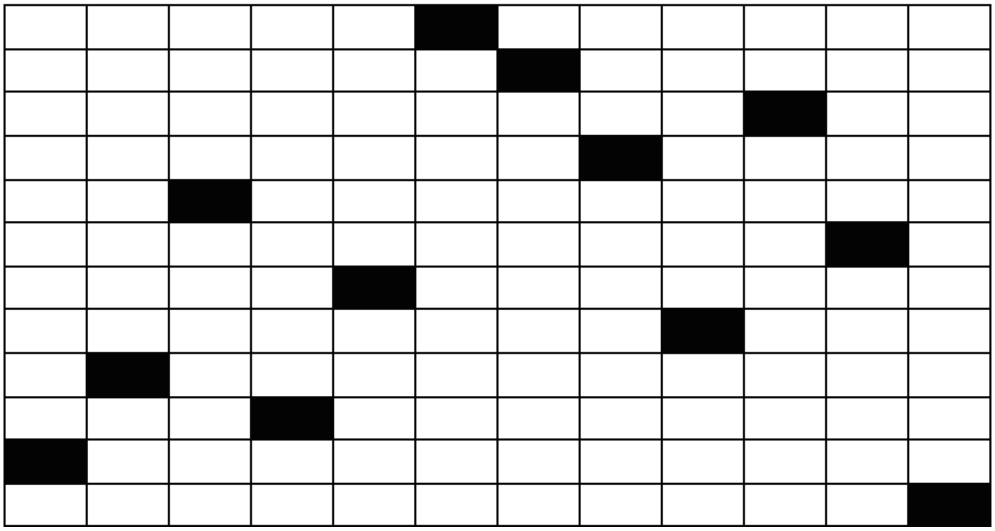
First, let's list some examples of beautiful intervals and describe the mathematical relationships between the notes that compose these. If an interval is an octave, then one note has double the frequency of the other; the frequencies of perfect fifths are in the ratio 3/2, perfect fourths in the ratio 4/3.

Now that we have some examples to extrapolate from, the next question becomes: Which ratios sound beautiful, and which inharmonious? Our instincts may tell us that all ratios with rational numbers are harmonious, but this isn't quite the case. An interval whose notes have a ratio of 2309⁄3000 is definitely not harmonious, but yet the ratio of the frequencies of the notes is rational. The reason why the above example does not result in a beautiful interval is that the pattern of the two overlapping frequencies is too complex for the human ear, and as a result, it does not sound pleasant. Conversely, there are intervals whose frequency ratios are irrational numbers that are very close to simple rational numbers and do actually sound pleasant. The range of such irrationals is very small. However, we shall assume that all rational numbers are harmonious as well as all irrational numbers extremely close to rational numbers.

Now, since irrational numbers greatly outnumber rational numbers, the number of cacophonous intervals greatly outnumbers the number of harmonious intervals. This bears a rather strange resemblance to a branch of mathematics known as measure theory, the study of sizes in mathematics. So let us use this connection. Suppose we measure the number line from 1 to 2. We want to create collections—open sets—that encompass all of the harmonious ratios resulting in beautiful intervals. So what do we do? We could just have one big open set going from 1 to 2, but that would include all of the ratios that describe unpleasant intervals. Thus, what we want to create is an infinite collection of numbers that neglects certain groups of numbers.

We can start by lining up everything that we know is harmonious: every single rational number. Let us order these simplified numbers in order of increasing denominator such that each appears only once: 3⁄2,4⁄3,5⁄3,5⁄4,7⁄4, and so forth. We know from geometric series that $1/2+1/2^2 +1/2^3 + \ldots =1$. If we were to represent each of our harmonious ratios with a distinct element in this series, the sum of our open sets would be at most 1. However, our open sets could still contain all of our inharmonious intervals. What we do then is create sets that contain as many or as few of the inharmonious intervals as we want by multiplying this equality by some n where $0 < n < 1$ to have sets of sizes n/2, n/4, n/8, and so forth. For instance, it might be reasonable to say that $n = .01$, and so only 1 percent of all possible intervals would be defined as "harmonious."

While harmony is certainly one aspect of beauty in music, another equally important one is melody. Harmony is the term used



12x12 Costas array. *Sam Markowitz.*

to describe notes played simultaneously in a piece of music, but melody describes the sequence of notes. So what makes a melody beautiful? One factor is the creation of a pattern, or motif. These themes are created, utilized, and explored throughout a piece of music. If beauty is a result of these patterns, then a piece of music without pattern would be considered ugly. In order to learn more about what makes a melody beautiful, we may find ourselves asking: How can we create the ugliest piece of music possible?

The goal here is to make no part of the song similar to another. Therefore, no notes are repeated. However, this also means that the melody cannot move in the same direction for the same amount more than once. For instance, if there is a note that moves up a whole step at the beginning of the piece, then we must make sure that no other note moves up a whole step for the remainder of the piece.

Before I show how to do this, let's do an experiment. Pick a prime number p. It can be any prime you want (though I recommend that it be relatively small). I will use 5 as an example. Then try taking any number, excluding 1, that is less than that prime and find its first p-1 powers. We will use the number 3, whose first 4 powers are 3, 9, 27, 81. Divide each of these numbers by your prime and leave only the remainders.

In this case, we will divide by 5, and our remainder sequence is 3, 4, 2, 1. Notice that no two adjacent notes move in the same direction by the same amount as any other pair of adjacent notes. I would hazard a guess that your numbers worked in exactly the same fashion. That is because we have created something called a Costas array. The amazing thing about these Costas arrays are that we use each number only once, and the difference between any two adjacent numbers is unique.

We can apply these numbers to notes in regular intervals. For instance, a difference of 1 in numbers may indicate a half step in music. By applying these arrays to music, we can create a truly ugly melody. Lecturer Scott Rickard gave a TED Talk discussing this very topic in addition to demonstrating his own piece of music based on a size 88 Costas array. Rickard also used something known as a Golomb ruler to create irregularity in the rhythm. This is an integer line with marks at certain positions such that the distance between any two marks is unique, whether or not the two are adjacent.

Thus, through this brief exploration of the relationship between music and mathematics, we are able to pinpoint some of the factors that make music beautiful. And to think! This is only a brief introduction to the musical applications of pure mathematics. We can also use such mathematical concepts to create and observe many other fascinating occurrences in music. •

BIBLIOGRAPHY

1. Costas array. Wikipedia. 2016 Nov 10 [accessed 2017 Feb 17].https://en.m.wikipedia.org/wiki/Costas_array
2. TEDT. YouTube. YouTube. 2011 Oct 3 [accessed 2017 Feb 17]. https://www.youtube.com/watch?v=RENk9PK06AQ&feature=youtu.be
3. YouTube. YouTube. 2015 Oct 3 [accessed 2017 Feb 17]. https://www.youtube.com/watch?v=cyW5z-M2yzw&feature=youtu.be

THOUGHT

EXPERIMENT

We Were in Screaming Color (But Can Colors Speak?)
Frank Jackson's Knowledge Argument

Mary Atwater knows everything there is to know about the color blue, and she has dedicated her life as a scientist to understanding everything from why the sky is blue, why Choate's school colors are blue and yellow, or why blue often represents tranquility. This would have been pretty unremarkable—everyone has a niche interest or two, right?—but Mary is completely color blind. She can only see it in black and white.

However, there was recent innovation in neurosurgery that could repair the broken transmission of signals from the cones on her retina to her brain. Mary will soon be able to see the world in stunning color.

Will she find something knew about the color blue? Or will she find that witnessing what it looks like adds nothing to her existing knowledge?

Submit your thoughts to tlee17@choate.edu for a chance to be featured in the next issue!

Selected Answer from Last Issue

Life is but a Dream (so should you row the boat?): Robert Nozick's Experience Machine

It's the news of the decade. Scientists have invented a technology called the Experience Machine that would allow you to experience your ideal life. The machine is guaranteed to work flawlessly. In order to sign up to use this machine, you need to go to a lab. You'll be placed in an eternal coma. You'll never wake up from the coma and interact with the real world ever again, but you are living out your wildest dreams. Would you sign up to use this machine?

No. The Experience Machine, as real as it would feel to someone hooked up to it, would still be fake. The idea of synthetic experience isn't what puts me off, rather it's that from the outside looking in, you would know that no matter what the person attached to it felt, it would all be fake.

Personally, I draw a lot of parallels to the old saying, "ignorance is bliss." Would you rather know and be miserable? Or would you instead choose to remain unknowing?

The Experience Machine is ignorance. To choose the Experience Machine is to choose to remain in the dark. While blissful, it's still ignorance—it's still fake. As a person who would be offered this choice, I would never be able to choose an eternal coma and a life of limitless make-believe.

Despite the fact that, were I to use the Experience Machine, I would neither know about the choice I'd made nor that the world I was experiencing was fake, I would find myself unable to choose willful ignorance.

Even if the truth hurts—even if reality hurts—it's real.

-Eilidh Dunsmore '19