CHOCOLATE

KEY

Dale J. Bachman, Ezra A. Brown,
and Anderson H. Norton

This colorful illustration of a primary component of modern cryptography—the Diffie-Hellman key exchange—draws students into the secret world of message encoding and decoding.

The word *cryptography* comes from the Greek *cryptos* (hidden) and *graphos* (writing), so cryptography is the science of hidden or secret writing. More generally, cryptography refers to the science of safeguarding information. If you have ever used a Web browser or a cell phone or bought anything over the Internet, you have used cryptography. And cryptography is becoming more and more important as we rely more and more on electronic—that is, digital—communication in business and in our personal lives.

Cryptography allows us to use a public medium such as the Internet to transmit private information securely, thus enabling a whole range of conveniences, from online shopping to personally printed movie tickets to (coming soon!) fraud-proof credit cards. To accomplish all this, cryptography must draw from the latest developments in mathematics and computer science, and its consequences define the cutting edge of privacy and intellectual property law.

From an education viewpoint, cryptography—as a result of its position at the intersection of mathematics and computer science, its implications for our daily lives, and its image as an element in

UTEHIL/ISTOCKPHOTO.COM

international espionage—is a rich field for spurring students to further study in the sciences. Although some of the smartest people in the world are involved in building and breaking new cryptographic algorithms, the basics of the subject and even some of the modern algorithms in common use are accessible to the average high school algebra student. To illustrate this last point, we study one of the most popular cryptographic algorithms today using M&M's® for a demonstration by analogy.

BACKGROUND

Modern encryption requires that the message sender and the message receiver share a common key, which determines the algorithms (keyed functions) used for encryption and decryption. To communicate over digital media, such as the Internet, we must convert messages into a string of zeros and ones. Encryption involves transforming that string into another, seemingly random, string of zeros and ones that no one can interpret without the key for determining the decryption algorithm. The key for encryption and decryption is also a string of zeros and ones, but we can think of that string as the binary expansion of an integer. For example, 11001 represents the integer 25.

Keyed functions allow us to exchange secret messages with our friends as long as we share a key with each friend. Such a system is called a *shared-key system*. The trouble with shared keys is that they have to be shared. If my friend lives halfway around the world, coming up with a key that both of us know but is secret from everyone else poses a real problem. It would be great if we could find a way to agree on a secret key (i.e., an integer) without exchanging secret information.

In the following activity, we will not study the actual encryption and decryption of messages. Rather, we focus on how two people might develop a secret key that would be used to produce encryptions and decryptions.

JO ANN SNOVER/ISTOCKPHOTO.COM

THE PROBLEM

The following analogy shows how we might come up with a public method of agreeing on a secret key. We have conducted this activity with audiences of all levels—from elementary school students to national decision makers—and all seem to understand and enjoy it (probably because we always let them eat the keys afterward!).

The setup includes a pair of empty cans with plastic lids (cans for nuts work well). In the lid of each can, cut a pair of crossed slits about a half-inch long; the slits allow the insertion of small objects into the cans without opening the lids but do not permit the can's contents to be seen. Choose two students on opposite sides of the class, give each a can, and supply each with a small pile of plain M&M's of many colors. For the duration of the demonstration, the students have to change their names to Alice and Bob (in cryptography, two people trying to communicate are always named Alice and Bob).

The cans will represent keys for encoding and decoding messages. We imagine that everyone in the class has an encryption-decryption machine that uses cans of M&M's as keys. That is, the machine is some kind of box that has a slot in the shape of a can, an input chute, an output chute, and a crank that can go forward or backward. If Alice wants to send a message to Bob, she picks a key (a can of M&M's), puts the key in the slot, puts her message in the input chute, and cranks forward so that an encrypted message comes out. She then sends the message to Bob. Even when she sends the message through a public network (such as the Internet), the message will remain secure unless someone else has the same key for encrypting-decrypting.

When Bob receives the encrypted message, he will place his identical key (a can with the same numbers and colors of M&M's as Alice has) in the slot of his encrypting-decrypting box. He will then put the encrypted message into the input chute and crank backward to get a decrypted message out. The tricky part is for Alice and Bob to obtain identical cans to begin with, without anyone else obtaining copies of those cans or guessing their contents (knowledge of the contents would allow someone else to create an identical can).

The rules of the game are as follows: Alice and Bob insert as many M&M's of whatever colors they like into their respective cans. No one can open either can, not even Alice or Bob. However, any student holding a can is able to duplicate it magically, including its contents. In practice, this means that each student who has touched the can pretends that he or she still holds an identical copy, even after passing along the original can (but still one cannot open the copy to see what is inside). Students can also insert M&M's into the copy. We assume that everyone has an encryption-decryption machine that uses cans of M&M's as keys; thus, if two students have cans with identical contents, they can encrypt and decrypt messages to each other. If anyone else also has an identical can, then he or she can read all the messages encrypted by using that can as the key. The goal is for Alice and Bob—and no one else—to end up with cans with identical contents, but all their messages to each other must pass through the classroom so that every student can touch and copy them. How can Alice and Bob achieve this goal?

A SOLUTION

Given enough time, most classes develop this solution themselves and often come up with some nice subtleties and side discussions. Allow students to investigate the concepts and arrive at a protocol, such as the following one.

Alice and Bob each put some M&M's in his or her can and remember which M&M's they used. Then they exchange the cans by handing them from student to student through the classroom (note that many students will have a chance to handle and make copies of the cans but cannot see the contents of any can). Once Alice and Bob receive the other's can, they each put the same M&M's as before into this second can. Now Alice has a can with Bob's M&M's and a copy of her own, and Bob has a can with Alice's M&M's and a copy of his own. In other words, they have cans with identical contents.

What about the other students who handled the cans? Many of them might have handled both cans, so under the rules of the game they could have copies of both cans. Could an eavesdropper use the copies of Alice and Bob's original cans to produce a copy of the can that Alice and Bob use as their key? This question leads to the conclusion that the eavesdropper would really like to open one of the cans. Then he or she would know everything that Alice (or Bob) knows and could produce the shared key.

THE ANALOGY: THE DIFFIE-HELLMAN KEY EXCHANGE

Let's do the same thing, replacing M&M's and nut cans with numbers and functions, respectively. The algorithm we describe is called the Diffie-Hellman key exchange, and it is a part of many of the communications systems we use every day.

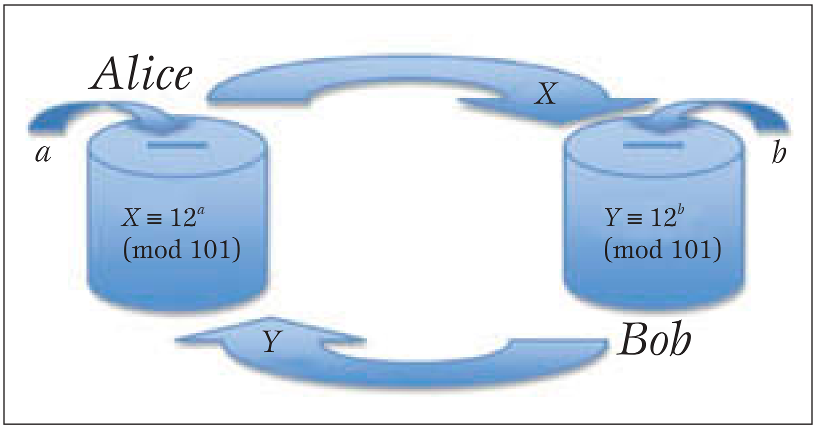
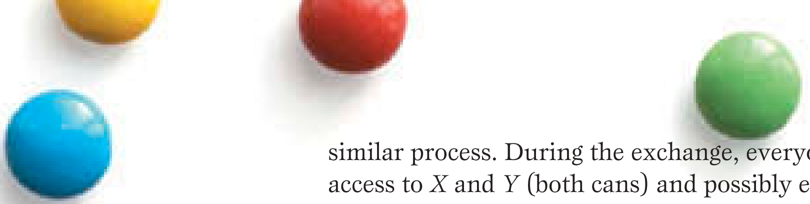


Fig. 1 The group (Z^*_{101}), the base (12), the value of X, and the value of Y may be known to all.

The context in which we will work is the algebraic structure called a *group*. A group is a set of elements and an operation such that the set is closed under the operation, the elements obey the associative property with respect to the operation, there is an identity element in the group under the operation, and each element in the group has an inverse with respect to the operation. Any group will work, but let's take a concrete example. We will use the numbers $\{1, 2, ..., 100\}$ under the operation multiplication modulo 101—that is, after we multiply two numbers, we divide by 101 and use the remainder as our answer. For example, $50 \cdot 3 = 150$, but because 150 is not one of the numbers in our group, we divide 150 by 101 and get a remainder of 49. So we say $50 \cdot 3 \equiv 49$ modulo 101 (i.e., $50 \cdot 3$ is "congruent" to 49, modulo 101). We call this group Z^*_{101}.

Next, we choose a base element from the group—say, 12. We will work with the powers of our base element. Now, Alice and Bob need to agree on one of the numbers in the group Z^*_{101} to use as a shared key. They start by each choosing an exponent, as represented by choosing their respective sets of M&M's. Let's say that Alice chooses 4 and Bob chooses 23. Then they compute their powers of 12, as represented by putting their collections of M&M's into their respective cans. Alice computes $12^4 \equiv 31$ (modulo 101), and Bob computes $12^{23} \equiv 35$ (modulo 101). Then they send these numbers to each other, as represented by exchanging cans; Alice receives 35 from Bob, and Bob receives 31 from Alice. Alice raises the number she received to the power that she knows, as represented by adding her M&M's to the can Bob gave her. She computes $35^4 \equiv 68$ (modulo 101), and similarly Bob computes $31^{23} \equiv 68$. Voilà, they get the same number, and they can use it as a shared key.

Figure 1 illustrates the process and analogy. Alice puts some collection of M&M's (integer a) into the can (function $X \equiv 12^a$ modulo 101) and exchanges cans with Bob, who has completed a



similar process. During the exchange, everyone has access to X and Y (both cans) and possibly even the base and group used in the functions but not to a and b. That knowledge would require opening the cans. Then Alice puts the same collection of M&M's in Bob's can that she put into her own (raising Y to the power of a). Likewise, Bob puts his same collection into Alice's can (raising X to the power of b). Both processes yield the same result.

How secure is this? What about spies? A third person (named Charlie, of course) may see both the 31 that Alice sent to Bob and the 35 that Bob sent to Alice, and he may also know that they are working in Z^*_{101} with 12 as their base element. He would love to know either Alice's or Bob's exponent because then he could compute their shared key. The problem of finding the exponent in this situation is called the discrete logarithm problem.

How might Charlie find one of the exponents without opening either can? One way would be to start listing the powers of 12 (modulo 101). Because there are only 100 numbers in our group, we might expect that after about 100 tries he would have exhausted all the possibilities and would be bound to find one or both exponents. However, Fermat's little theorem guarantees that $12^{100} \equiv 1$, so if Charlie tries exponents larger than 100, he will find that he is repeating himself. In fact, he only has to list a few powers of 12 ($12^1 \equiv 12$, $12^2 \equiv 43$, $12^3 \equiv 11$, $12^4 \equiv 31$) before he notices the number Alice sent to Bob and determines that her exponent must be 4.

Thus, we choose a very large group (and very large exponents) so that it is hard for Charlie to list powers and find our exponents. There are ways of finding exponents that are more clever than listing powers. For example, experts say that if you want to be sure that no one reads your mail, you should choose a group of size at least 2^{1000}, which means that the numbers you are using (both elements and exponents) will be at least 1000 bits, or about 300 decimal digits, long.

MATHEMATICAL CONNECTIONS

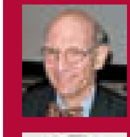
The cryptographic demonstration shared here introduces possibilities for several mathematical connections, some already included in the secondary school curriculum and others extending typical secondary school topics. The most obvious connection to the secondary school curriculum concerns exponents. Beyond computing exponents, investigations of the Diffie-Hellman key exchange invite students to consider logarithms for inverting exponential functions and properties of exponents that guarantee that Alice and Bob will arrive at the same result. And considering logarithms in a discrete context, such as working with the integers $\{1, 2, \ldots, 100\}$ in Z^*_{101}, leads to the discrete logarithm problem, which arises in college-level algebra and number theory courses (For more, see www.rsa.com/rsalabs/node.asp?id=2193). Likewise, the division algorithm used in modular arithmetic comes up in both high school and college-level algebra, and it generalizes to polynomial long division.

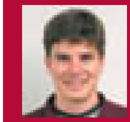
Encrypting and decrypting messages over digital media, such as the Internet, introduces the need for binary arithmetic because computers deal with only zeros and ones (switches that are either off or on). All messages are translated as zeros and ones, which keyed functions can encrypt and computers can send to one another. Further, generating keyed functions for encryptions that eavesdroppers cannot crack relies on the use of large groups of the form Z^*_p, where p is a very large prime number. As a result, developing computer programs that generate very large primes is a hot topic in number theory and computer science. In fact, UCLA's mathematics department recently earned acclaim for finding a thirteen-million-digit Mersenne prime (a prime number of the form $2^p - 1$ where p is itself a prime).

CONCLUSION

We recommend this cryptography demonstration and analogy as an engaging and colorful activity that invites students to consider topics at the cutting edge of mathematics and computer science. It is one classroom activity that will not be greeted with the ubiquitous "When are we ever going to use this?"

We all use cryptography every day, whether we realize it or not. Acknowledging this reality allows students to be more informed and more mathematically efficient.





DALE J. BACHMAN, dbachman@ucmo.edu, is an associate professor of mathematics at the University of Central Missouri in Warrensburg. In addition to cryptography and computer security, he researches algebra. EZRA A. "BUD" BROWN, ezbrown@vt.edu, is the Alumni Distinguished Professor of Mathematics at Virginia Polytechnic Institute and State University (Virginia Tech) in Blacksburg. His research interests include graph theory, combinatorics, and mathematical connections. ANDERSON H. NORTON, norton3@vt.edu, is an assistant professor in the mathematics department at Virginia Tech. He conducts research on students' mathematical development.