LOTTA WEDMAN

Kryptering på gymnasiet

Gymnasiets kurs Matematik – diskret innehåller flera av de ingredienser som behövs för att förstå och knäcka några av historens kända krypteringsmetoder. Genom att arbeta med kryptering får eleverna inte bara möjlighet att öva sina matematikkunskaper utan får också se att kursens matematikinnehåll är praktiskt användbart.

Bakgrunden till att jag använde kryptering i kursen *Matematik – diskret* är att jag sökt efter tillämpningar av matematik på gymnasienivå. Vårterminen 2003 var första gången som kursen, som är obligatorisk på naturvetenskapliga programmets datainriktning, hölls. Parallellt med kursen läste jag *Kodboken* av Simon Singh. Jag prövade att efter kombinatorikavsnittet lära eleverna lite kryptering (motsvarande fyra lektioner), dels som ett tillämpningsområde i kombinatorik och dels som en övergång till talteorin. Eleverna var positiva. När jag läste vidare fann jag mer och mer matematik och nästa gång kursen hölls utgick vi från krypteringen. Jag ville se hur stor del av kursen man kunde få in, helst utan att tappa tempo. Jag hade den våren fyra elever och vi höll på med kryptering ungefär en tredjedel av kursen. Under den tiden uppnådde vi följande kursmål:

Eleven ska

- kunna formulera, analysera och lösa matematiska problem av betydelse för tillämpningar och vald studieinriktning med fördjupad kunskap om sådana begrepp och metoder som ingår i tidigare kurser,
- vara förtrogen med viktiga egenskaper hos mängden av de hela talen och kunna tillämpa dessa i problemlösning,
- kunna använda grundläggande begrepp och principer inom kombinatorik samt tillämpa dessa vid t ex analys av algoritmer.

Dessutom fick vi med:

- mer talteori än vad jag tog upp första gången jag hade kursen,
- datorhistoria,
- andra världskrigets historia,
- grammatik,
- större förståelse för vad matematiker kan arbeta med.

Vi började kursen med en introduktion till kryptering. Vi gick bland annat igenom hur kryptering har använts genom historien och hur man använder den idag. Sedan tog vi upp några olika krypteringsmetoder. Först lärde vi oss förstå och använda metoderna och sedan diskuterade vi hur man forcerar de krypterade meddelandena. När vi kommit in på den matematik som ingår i kursen tog vi fram boken och räknade.

Lotta Wedman *är lärare i matematik och filosofi vid Västerbergslagens utbildningscentrum i Ludvika.*
lotta.wedman@vbu.ludvika.se.

Kursens upplägg

Vårt upplägg i kursen var i stort sett kronologiskt och därför kommer jag här att beskriva de stora dragen i krypteringens historia och vilken typ av matematik som vi stötte på när vi jobbade. Först ett litet exempel på hur kryptering kan gå till som förklaring av begreppen:

> En sändare vill skicka ett meddelande till en mottagare utan att någon obehörig kan tyda det. Hon krypterar först klartextmeddelandet, d v s meddelandet i läsbart skick, med hjälp av en algoritm och en nyckel, som specificerar algoritmen. Kryptotexten skickas till mottagaren som, då han känner till algoritmen och nyckeln, förvandlar texten tillbaka till läsbart skick. Obehöriga kan snappa upp den krypterade texten men ska inte kunna forcera kryptot och tyda texten.

En av de första uppgifterna eleverna fick såg ut så här: ”Hitta det kryptoalfabet med vilket denna text är krypterad.” Vi hade då gått igenom *monoalfabetiska substitutionskrypton* i vilka varje tecken i klartextalfabetet byts ut mot ett och endast ett tecken i kryptotextalfabetet. Ett exempel på en sådan kryptering är att förskjuta alfabetet ett antal steg. De krypton som brukar förekomma i dagstidningar är också monalfabetiska. Dessa kryptotexter är enkla att forcera med hjälp av logik och kunskap om det svenska språket.

Eleverna fick till sin hjälp en tabell över relativa frekvenser för våra bokstäver. ”E” är vanligast (9,9 %) följt av ”a” (9,3 %) ”n” (8,8 %) och ”t” (8,7 %). Vi satt någon timma och räknade på frekvenser och klurade. Även om vi får fram att ”p” är vanligast i vår kryptotext kan vi inte förutsätta att det är ett ”e” i klartext. Det skulle kunna vara ett ”a”, ”t” eller kanske någon annan bokstav beroende på vilken text det handlar om. Det är ändå troligt att det är en av de vanligaste bokstäverna. Vissa ord är lätta att känna igen. ”Och” är ett vanligt ord med relativt ovanliga bokstäver i. ”I” är ett annat exempel. Några elever struntade helt i frekvensberäkning och prövade sig fram istället. De flesta knäckte meddelandet utan större svårigheter och vi insåg att det krävs lite mer avancerade krypteringssystem.

Vigenèrekryptering är ett exempel på ett *polyalfabetiskt* krypto. Man använder sig av flera olika kryptoalfabet där varje alfabet är en förskjutning. För att veta vilka alfabet man ska använda har man ett kodord. Om kodordet är GRODA använder man sig först av det krypteringsalfabet där bokstaven A blir ett G. När man krypterat den första bokstaven byter man alfabet till det där A blir R och krypterar andra bokstaven. När man krypterat sig igenom fem bokstäver så börjar man om igen och den sjätte bokstaven krypteras med hjälp av G-alfabetet igen.

Första uppgiften på Vigenèrekrypton var:

> a) Välj en mening, vilken som helst, och kryptera den med hjälp av en Vigenèretabell och kodordet ”GRODA”.
> b) Lämna över din krypterade text till de andra i gruppen.
> c) Återomvandla de andras meningar till klartext.

Det är lite knepigt att kryptera meddelanden för hand utan att det blir fel och framförallt att läsa dem. Uppgiften var svårare än den verkade. Vi insåg hur skönt det är att låta datorer göra den här typen av arbeten. Fördelen med den här typen av krypton jämfört med de monoalfabetiska är att det inte alls är lika lätt att forcera systemen. Vigenèrekryptot kallades länge för ”Le Chiffre Indéchiffrable”.

Talteori till hjälp

När vi diskuterade hur ett krypterat meddelande skulle forceras kom vi in på mycket av den talteori vi tog upp i kursen. Först letar vi i kryptotexten, som bör vara ganska lång, efter sekvenser som återkommer. Kanske hittar vi ”EFIQ” med 95 bokstävers mellanrum. Dessa sekvenser kan visserligen ha uppstått genom en slump men vi utgår från att de har krypterats med nyckeln på samma sätt och att sekvenserna står för lika sekvenser i klartextmeddelandet. För att minska risken för det första alternativet håller vi oss till sekvenser på minst fyra bokstäver. Vi utgår sedan från avståndet mellan de sekvenserna. Om avståndet mellan ”EFIQ” är 95 bokstäver, mellan ”MMNA” 20 och mellan ”KPIDE” 120, borde kodordets längd

vara en faktor i både 95, 20 och 120. Vi faktoruppdelar de tre talen; $95=5\cdot19$, $20=2^2\cdot5$ och $120=2^3\cdot3\cdot5$. De gemensamma delarna är 1 och 5. Kodordet är i det här fallet antingen 1 bokstav långt eller 5 bokstäver långt. Att kodordet är 1 bokstav långt har vi antagligen redan förkastat eftersom meddelandet i så fall skulle vara krypterat med ett monoalfabetiskt krypto. Vi antar därför att det är 5 bokstäver långt. Sen ska texten delas upp i fem olika deltexter (var femte bokstav i samma text) och dessa forceras med hjälp av frekvensanalys. Analysen underlättas av att vi vet att alfabeten är förskjutningar och att kodordet ofta är ett ord. Vi gick inte så långt att vi forcerade meddelandet helt. Vi nöjde oss med att konstatera hur långt kodordet var. Om vi skulle ha gått vidare borde vi ha använt oss av datorer.

Enigma

Eftersom Vigenèrekrypton går att knäcka med lite vilja om kodordet är relativt kort blir det naturligt att fundera över system med riktigt långa kodord. Ett sådant system är krypteringsmaskinen Enigma. Vi ägnade ganska mycket tid till att förstå hur Enigma fungerar. På ett förslag från en av eleverna gjorde vi en pappersmodell av en liten maskin som vi krypterade med. Enigmas viktigaste beståndsdel är en så kallad slumpkodare som består av tjocka gummiskivor fulla av elledningar. En slumpkodare tilldelar varje bokstav i klartextmeddelandet en krypterad bokstav. När den har krypterat en bokstav, snurrar den ett steg och byter kryptoalfabet.

Man kan sätta flera slumpkodare efter varandra så att när den första har snurrat ett varv snurrar den andra ett steg och när den andra har snurrat ett varv så snurrar den tredje ett steg och så vidare. Om man till exempel har tre olika slumpkodare som kan snurra runt ger dessa upphov till 26^3 olika kryptoalfabet (givet att vi har ett alfabet med 26 tecken). Har vi fem olika slumpkodare av vilka vi ska välja ut tre i en viss ordning så ger detta upphov till $5\cdot4\cdot3\cdot26^3$ olika alfabet. Sen ökar vi antalet ytterligare genom att lägga till en kopplingstavla där vi genom kablar kan byta plats på bokstäver parvis. Det blir betydelsefullt att bygga en maskin med så många olika alfabet som möjligt för att försvåra forceringen. Mellan lektionerna var eleverna ute på Internet och hittade sidor där vi kunde testa att kryptera texter med hjälp av en Enigmamaskin. Vi kunde välja vilka slumpkodare som skulle användas och hur kablarna skulle ställas in. Se www.ugrad.cs.jhu.edu/~russell/classes/enigma/

När vi arbetade med Enigma fann vi tillämpningar av nästan hela den kombinatorik vi tog upp i kursen. Efter denna inledning plockade vi fram läroböckerna i matematik. Men vi kom tillbaka till krypteringen. Jag höll en föreläsning om hur Enigma knäcktes och vi kom in på mer moderna krypteringsmetoder.

Datorkryptering har färre begränsningar, är snabbare och använder sig av siffror istället för bokstäver. Alla meddelanden omvandlas till binära tal innan man krypterar dem. ASCII-kod tilldelar varje bokstav i alfabetet ett binärt tal. Det finns 2^7 olika sätt att ordna sju binära siffror på och därför kan ASCII-kod beteckna upp till 128 olika tecken. När väl konverteringen är genomförd kan man kryptera texten. Detta var inget vi arbetade med mer än rent teoretiskt eftersom det är ett område som jag inte behärskar.

Öppen nyckel

Ett problem vid kryptering är frågan om nyckeldistributionen. Hur ska man överföra nyckeln på ett säkert sätt? RSA-kryptering är ett system som inte behöver någon hemlig nyckel, eftersom det bygger på något som kallas en envägsfunktion. RSA står för Rivest, Shamir och Adleman vilka var namnen på de som skapade metoden.

All kryptering bygger på en funktion som har en invers. Funktionen *addera 2* har tex inversen *subtrahera 2*. De flesta matematiska funktioner som har en invers är tvåvägsfunktioner, dvs om vi känner till funktionen så kan vi lätt hitta dess invers. En dubblering är ett exempel på en tvåvägsfunktion. Om vi vet att resultatet av en dubblering är 18, är det lätt att kasta om och härleda det ursprungliga talet, nämligen 9.

En envägsfunktion däremot, är en funktion för vilken det är väldigt svårt att hitta inversen även om vi känner till funktionen.

Inom området modulär aritmetik finns flera exempel på envägsfunktioner. En av dem ligger till grund för RSA-krypteringen.

RSA-kryptering går till så att mottagaren väljer två stora primtal p och q. Ett känt resultat inom talteorin garanterar att det är ganska lätt att hitta även mycket stora primtal med hjälp av en dator. Mottagaren multiplicerar sedan sina två primtal och får $N=p\cdot q$. Han väljer också ett tal e så att e och $(p-1)\cdot(q-1)$ blir relativt prima; största gemensamma delaren är 1. N och e blir mottagarens offentliga nyckel som han publicerar. Talen p och q håller han hemliga.

När en sändare vill kryptera ett brev till mottagaren letar hon rätt på hans publika nyckel (N,e). Hon konverterar sitt meddelande till ASCII-kod, kalla meddelandet i denna form M, och krypterar M med hjälp av funktionen $C=M^e \pmod N$. Resultatet C, det krypterade meddelandet, skickar hon över till mottagaren.

När mottagaren ska dekryptera meddelandet, nu i form av talet C, räknar han först ut talet d så att $e\cdot d=1 \pmod{(p-1)\cdot(q-1)}$ med hjälp av *Euklides algoritm*. Det visar sig att med detta val av d så gäller $X^{de}=X \pmod N$. Mottagaren kan då få fram originalmeddelandet genom $M=C^d \pmod N$ eftersom $C^d=(M^e)^d=M^{ed}=M \pmod N$. Slutligen konverterar mottagaren meddelandet tillbaka till text.

Krypteringfunktionen ovan är exempel på vad som kallas en envägsfunktion. Med vetskap om N och e kan vi lätt beräkna funktionen. Utan mer exakt information så är det dock mycket svårt att hitta funktionens invers som vi behöver för att dekryptera meddelandet. För att hitta inversen måste man nämligen känna till primtalen p och q, inte bara deras produkt N. Om primtalen är små kan man lätt hitta dem när man vet N, men när vi valt dem tillräckligt stora så är det mycket, mycket tidskrävande att faktorisera N för att få reda på p och q.

För att förstå hur RSA-kryptering går till måste man behärska stora delar av gymnasiets talteori; primtal, binära tal, kongruensräkning och Euklides algoritm. Om man dessutom vill förstå hur systemet är uppbyggt måste man in på mer avancerad talteori, det bygger bland annat på Eulers ϕ-funktion. Eleverna insåg att om man ska arbeta med kryptering måste man dels kunna en del matematik, mer än vad man lär sig på gymnasiet, och dels kunna programmera datorer.

Som en avslutning på temat diskuterade vi kvantdatorer och kvantkryptering som en möjlig framtid respektive ett helt säkert krypteringssystem. I takt med att datorerna blir snabbare blir det lättare att faktorisera stora tal och RSA-kryptering som är säkert idag kanske inte är säkert imorgon.

I provet, liksom i undervisningen, varvades kryptering med ren matematik. Exempel på uppgifter:

1. Tolka följande meddelande med hjälp av kodordet MJÖLK och en Vigenèretabell.

   EJCSOCNQOB

2. Hur knäcker man ett Vigenèrekrypto? Varför går det inte att göra så med ett meddelande krypterat med hjälp av Enigma?

3. Primtalsuppdela talet 180.

4. Hur vet man, utan miniräknare, om ett tal är delbart med
   a) 2 b) 3 c) 5

5. Använd Erathostenes såll för att hitta alla primtal mindre än 30.

6. Av 7 olika slumpkodare ska du plocka ut 3 och sätta ihop dem i ENIGMA. På hur många olika sätt kan du göra detta?

7. I det svenska alfabetet finns 29 olika bokstäver. Hur många olika kryptoalfabet ger 3 slumpkodare?

Området kryptering tog ungefär en tredjedel av kurstiden. Jag ville med krypteringen visa eleverna att diskret matematik går att använda. Trots att vi ägnade så mycket tid till annat än att räkna hade vi inte svårare att hinna med kursmålen jämfört med året innan.

Jag vill inte dra några slutsatser utifrån elevernas utvärderingar eftersom jag denna vår endast hade fyra elever men jag som lärare tyckte att undervisningen på det här sättet blev roligare och mer motiverande. Om jag skulle fortsätta utveckla projektet skulle vi nästa gång använda oss mer av datorer. Det vore också intressant om man kunde använda sig av liknande projekt i andra matematikkurser.