KODER I KLASSEROMMET

Kristian Ranestad

28.02.2001

Dette heftet er utarbeidet til klasseromsprosjektet ved Matematisk institutt, UiO. I dette prosjektet inngår det halvdags kurs for lærere i forskjellige tema. Dette heftet er ment som et kurshefte til et slikt kurs i hemmelige koder (kryptologi) og feilrettingskoder. Samtidig kan det forhåpentligvis også fungere som et ressurshefte for de som vil arbeide med disse emnene i sin undervisning, for eksempel i form av prosjektarbeid.
Innholdet er delt mellom praktiske eksempler og teori. De første kapitlene tar for seg enkle og illustrerende eksempler på bruk av hemmelige koder og feilrettingskoder. Det neste kapittelet tar for seg kongruensregning og løsning av lineære kongruenser, mens de siste kapitlene bruker denne teorien til å lage koder. I kapittel 6 og 7 har vi samlet ideer til undervisningsopplegg og til prosjektoppgaver beregnet for elever i videregående skole.
Til slutt har jeg med tillatelse fra forfatteren lagt ved Ben Johnsens artikkel "Kryptografi- en gammel disiplin med moderne anvendelser". Denne gir en fin historisk oversikt over bruk av kryptografi.

Oslo, 2001
Kristian Ranestad

Innhold

Tillegg:

Ben Johnsen: Kryptografi- en gammel disiplin med moderne anvendelser. (fra Per Hag & Ben Johnsen (red.): *Fra Matematikkens Spennende Verden*, Tapir, 1993, 123-134.)

1. Hemmelige meldinger

Dersom Ole vil sende en melding til Kari uten å røpe innholdet i meldingen til andre, kan det være mange måter å gjøre dette på. Hvis de er redd for at meldingen blir snappet opp av andre på veien mellom de to kan de kode meldingen, omforme den på en måte som de selv kjenner, men som eventuelle utenforstående ikke kjenner. Meldingen er da kryptert og er blitt hemmelig. Selv om det kan høres litt spesielt ut er det i prinsippet det samme med en kodelås for en sykkel. Låsen og eieren kjenner koden som skal til for å åpne låsen, men den er hemmelig for utenforstående. Vi skal i hovedsak se på ulike metoder til å gjøre tekstmeldinger om til hemmelige meldinger. I artikkelen til Ben Johnsen [1] er det en fin historisk oversikt over bruken av slike hemmelige koder.

1.1 Eksempel

Vi går tilbake til Ole og Kari. Her er en kodet melding som Ole sender til Kari:

J MPWF V

og her er samme melding kodet på en annen måte:

U EVOL I

Hvis en får litt tid klarer en å finne den opprinnelige meldingen, nemlig

I LOVE U

I begge tilfeller er ikke den hemmelige meldingen særlig hemmelig, siden den lar seg tyde uten for store vanskeligheter. Men de to eksemplene viser to prinsipper som de fleste koder bygger på.

Det første prinsippet kalles **substitusjon**, og går ut på at en erstatter bokstavene eller tegnene i meldingen med andre bokstaver eller tegn. I eksempelet er hver bokstav i meldingen erstattet med den bokstaven som kommer etter i alfabetet. Det er mange måter å substituere på. Vi skal i hovedsak erstatte bokstaver med tall, men her er det mange muligheter, og med en kort melding vil de fleste substitusjoner være vanskelige å tyde for utenforstående. Hemmeligheten ligger jo i at den utenforstående må finne ut på hvilken måte substitusjonen foregår, etter hvilken oppskrift så og si. I eksempelet er oppskriften ikke så vanskelig å komme på, men hvis substitusjonen erstatter bokstavene **E,I,L,O,U,V** med tegnene **%,+,=,*,@,&** ville den kodete
meldingen se slik ut

+=*&%@

og så ville det være umulig å tyde meldingen, hvis en ikke visste noe om substitutsjonen.

Det andre prinsippet kalles ombytting eller **permutasjon**, og går ut på å bytte om på

rekkefølgen av de bokstavene eller tegnene som er brukt i meldingen. I eksempelet er rekkefølgen rett og slett byttet helt om, slik at en får den kodete meldngen ved å lese den opprinnelige baklengs. Det fins selvfølgelig andre permutasjoner også, og siden meldingen er kort så er det et ikke uoverkommelig arbeid å undersøke alle permutasjoner for å tyde den hemmelige meldingen. En permutasjon av meldingen i eksempelet vil gi følgende hemmelige melding:

E VULI O

Denne er vanskeligere å tyde en baklengskoden, men med litt tid kunne en nok klare å tyde denne også (hvis en visste at det var en permutasjon som var brukt).

I praksis bruker man ofte både substitusjon og permutasjon for å lage hemmelige meldinger. Substitusjonen er da ofte fra bokstaver og tegn til tall, mens permutasjonen er en ombytting av tall.

For enkelhets skyld bruker vi et alfabet med bare 6 bokstaver. Det norske alfabetet har 29 bokstaver, og hvis en skal ta med tall og tegn blir det raskt over 50 ulike tegn som en eventuelt skal kode. Datamaskiner bruker den såkalte ASCII koden til å kode bokstaver og tegn til tall mellom 0 og 128 i det binære tallystemet. For eksempel kodes a til 1100001, A til 1000001, og tegnet & til 0100110. Disse blir så satt sammen i blokker av 8 slike tall. Dermed opererer man med et utvidet "alfabet" med opp til 2^{56} ulike blokker av tegn. Hver blokk er en sekvens av 56 nuller og enere, som så kan kodes videre.
Vi skal altså foreløpig holde oss til et alfabet med 6 bokstaver. I første omgang lar vi alfabetet være de bokstavene vi brukte i det første eksempelet. Satt i alfabetisk rekkefølge er det:

EILOUV

Først substituerer vi dem til tall, vi velger tallene fra 1 til 6 og setter substitusjonen opp slik:

E	**I**	**L**	**O**	**U**	**V**
↓	↓	↓	↓	↓	↓
1	2	3	4	5	6

Denne substitusjonstabellen kaller vi en **kodenøkkel**, fordi den forklarer eller avslører hvordan en får en kodet melding fra en melding i det opprinnelige alfabetet.

Meldingen

I LOVE U

blir med denne kodenøkkelen kodet til

234615

Uten kjennskap til alfabetet vårt ville selvsagt denne kodete meldingen være vanskelig om ikke umulig å dekode, altså å finne den opprinnelige meldingen til. For å bevare analogien

med koding i datamaskiner, vil vi imidlertid gå ut fra at denne kodenøkkelen er åpen, det vil si alminnelig tilgjengelig og kjent. Dermed er ikke den kodete meldingen mer hemmelig enn den opprinnelige. Den er bare litt mer tungvinn for oss å lese fordi vi må bruke kodenøkkelen for å oversette den til en forståelig melding.

For å gjøre meldingen hemmelig koder vi en gang til med en permutasjon. En permutasjon skriver vi også opp i en tabell. For eksempel er

1	2	3	4	5	6
↓	↓	↓	↓	↓	↓
6	4	1	3	2	5

permutasjonen som bytter 1 med 6, 2 med 4, o.s.v. Forsåvidt er dette også en substitusjon, men vi kaller det en permutasjon siden de tallene en substituerer til er de samme som en starter med. En slik permutasjon har en omvendt permutasjon, nemlig den som bytter tilbake. I eksempelet blir det permutasjonen

1	2	3	4	5	6
↓	↓	↓	↓	↓	↓
3	5	4	2	6	1

Når vi vil gjøre den kodete meldingen

234615

om til en hemmelig melding, kan vi bruke permutasjonen

1	2	3	4	5	6
↓	↓	↓	↓	↓	↓
6	4	1	3	2	5

som **hemmelig kodenøkkel** eller **krypteringsnøkkel**. Det vil si at vi
antar at denne kodenøkkelen bare er kjent for brukerne. Den hemmelig kodete eller krypterte meldingen blir da

413562

Uten å kjenne krypteringsnøkkelen blir denne meldingen uleselig selv om en kjenner til alfabetet (og den åpne kodenøkkelen).
For å finne tilbake til den opprinnelige meldingen må mottakeren av den krypterte meldingen først dekryptere med dekrypteringsnøkkelen

1	2	3	4	5	6
↓	↓	↓	↓	↓	↓
3	5	4	2	6	1

altså den omvendte av krypteringsnøkkelen. Den dekrypterte meldingen blir

234615

som selvfølgelig så dekodes og leses med den opprinnelige kodenøkkelen.

1.2 Prosedyre

Hele prosessen fra avsender til mottaker kan vi sette opp i et skjema:
MELDING:

I LOVE U

KODENØKKEL (ÅPEN):

E	**I**	**L**	**O**	**U**	**V**
↓	↓	↓	↓	↓	↓
1	2	3	4	5	6

KODET MELDING:

234615

KRYPTERINGSNØKKEL (HEMMELIG):

1	2	3	4	5	6
↓	↓	↓	↓	↓	↓
6	4	1	3	2	5

KRYPTERT MELDING:

413562

DEKRYPTERINGSNØKKEL (HEMMELIG):

1	2	3	4	5	6
↓	↓	↓	↓	↓	↓
3	5	4	2	6	1

DEKRYPTERT MELDING:

234615

DEKODENØKKEL (ÅPEN):

1	2	3	4	5	6
↓	↓	↓	↓	↓	↓
E	**I**	**L**	**O**	**U**	**V**

DEKODET MELDING:

I LOVE U

1.3 Eksempel

Vi kan på en enkel måte utvide alfabetet til alle bokstavene og fortsatt kode dem med tallene fra 1 til 6. Nå må vi imidlertid bruke to tall for hvert tegn. Vi bruker 2-sifrede tall i 7-tallsystemet som ikke inneholder 0. De første slike tallene er $11, 12, 13, 14, 15, 16, 21, 22, \ldots$. Kodenøkkelen ser da slik ut

A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	...	∅	**Å**
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	...	↓	↓
11	12	13	14	15	16	21	22	23	24	...	54	55

En melding som er kodet med denne kodenøkkelen kan så krypteres med en krypteringsnøkkel som er en permutasjon av tallene 1 til 6.
Dersom den kodete meldingen inneholder mange fler enn våre 6 tall, ville det vært vanskelig å huske en krypteringsnøkkel slik vi har laget her. Derfor er det naturlig å finne metoder til å lage gode permutasjoner som er lettere å huske, for eksempel ved at en ikke trenger å huske hele substitusjontabellen, men bare hvordan en finner de substitusjonene i en permutasjon som en til enhver tid har bruk for. Til dette kan vi bruke "moduloregning" eller kongruensregning. Vi skal gi en kort innføring i slik regning i kapittel 3 og så komme tilbake til hvordan vi kan lage krypteringsnøkler som er enkle å bruke.

I neste kapittel skal vi se på feilrettingskoder.

2. Feilrettingskoder

Feilrettingskoder er koder som kanskje er vel så mye i bruk som hemmelige koder. Bruken av slike koder er ofte litt skjult for oss, slik som for eksempel i signaloverføringer, både av blide og lyd. Når bilde eller lyd-data blir sendt over store avstander oppstår det lett fortyrrelser, støy eller feil i dataene. Disse dataene er imidlertid ofte kodet med en feilrettingskode. Da kan mottakeren korrigere feilene i de mottatte dataene, slik at vi får gode TV-bilder og god lyd likevel. En mer åpenlys bruk av feilrettingskoder kjenner vi fra strekkodene som blir brukt til å identifisere og prise varer i butikker.
En feilrettingskode er en kodenøkkel som er slik at det er mulig å oppdage feil i den kodete meldingen uten at en kjenner den opprinnelige meldingen. Feil kan oppstå når en sender en kodet melding, men som regel klarer en å redusere feilene til et minimum. Derfor er det nyttig for mottakeren å vite at dersom det bare er oppstått en feil i den kodete meldingen så er kodenøkkelen slik at hun kan oppdage at der er en feil, og muligens også rette denne feilen. La oss se på et eksempel.

2.1 Eksempel

Fra eksempel 1.1 bruker vi alfabetet **EILOUV**, men denne gangen bruker vi kodenøkkelen

E	**I**	**L**	**O**	**U**	**V**
↓	↓	↓	↓	↓	↓
11	22	33	44	55	66

Meldingen

I LOVE U

blir med denne nøkkelen kodet til

223344661155

Dersom det oppstår en feil, det vil si at et siffer blir forandret, før mottaker får den kodete meldingen, vil hun umiddelbart se at en feil er oppstått. Dette er selvsagt fordi hvert siffer gjentas en gang i den kodete meldingen. Vi sier at koden **oppdager** en feil. Dersom den mottatte meldingen er

223244661155

ser mottakeren at det er oppstått en feil i den andre bokstaven i meldingen (men hun kan ikke vite om den andre bokstaven skulle være kodet til 22 eller 33, altså om den andre bokstaven i meldingen var en **I** eller en **L**).
Dersom kodenøkkelen hadde repetert sifrene tre ganger

E	**I**	**L**	**O**	**U**	**V**
↓	↓	↓	↓	↓	↓
111	222	333	444	555	666

ville mottakeren kunne **oppdage og rette** den mottatt meldingen, dersom det bare hadde oppstått en feil. Slik kunne en selvsagt fortsette, men snart ville de kodete meldingene bli svært lange i forhold til den opprinnelige meldingen. Derfor har en bruk for gode kodenøkler som er slik at en kan oppdage og eventuelt rette enkle feil som er oppstått i kodete meldinger uten at meldingen blir uforholdsmessig store.

Strekkoden.

Strekkodene som er brukt i butikker er laget slik at dersom leseren i kassa leser en feil, så oppdager den det. Den får ikke kodene til å passe med listen av kodete varer. Dermed får ikke ekspeditøren registrert noen pris og prøver derfor en gang til. Det er så enkelt å prøve flere ganger, at det er tilstrekkelig at koden **oppdager** en feil. Strekkodene består av en rekke streker med forskjellig tykkelse. For at strekkodene skal kunne brukes til å identifisere varer, må de selvsagt være forskjellige. Hvis to strekkoder er forskjellige bare i en av strekene vil de to kodene lett kunne forveksles. Hvis leseren i kassa leser feil i en strek, ville det i så fall være umulig for ekspeditøren å vite om det var feil vare eller feil kode. Dersom minst to av strekkodene alltid er forskjellige i to streker, ville derimot en lesefeil bli registrert som feil kode (og ikke riktig kode for feil vare). Vi sier at strekkoden oppdager en feil. Derfor er to strekkoder alltid forkjellige på minst to plasser.

Senere skal vi se hvordan både ISBN-numre og fødselsnumre er kodet med en feilrettingskode.

3. Kongruensregning

I dette kapitlet lar vi t være et naturlig tall større enn 1. Dersom vi deler et helt tall med t, har vi t mulige rester: $0, 1, 2, \cdots, t-1$. Vi kan regne med tallene $0, 1, 2, \cdots, t-1$ ved bare å bruke disse restene. Det vil si at hver gang vi har to av disse tallene kan vi legge dem sammen eller multiplisere dem og få et nytt tall, nemlig ved først å addere eller multiplisere på vanlig måte og så ta resten etter divisjon med t. Denne regningen med tallene $0, 1, 2, \cdots, t-1$ kaller vi kongruensregning eller regning modulo t. To tall a og b har samme rest modulo t hvis og bare hvis differansen $a - b$ er delelig med t. I så fall sier vi at de to tallene a og b er kongruente modulo t og skriver $a \equiv b \mod t$. Derfor kalles også den regningen vi har beskrevet kongruensregning.

3.1 Eksempel (Regning modulo 2)

Hvis en legger sammen to partall eller to oddetall så får en et partall, og hvis en legger sammen et partall og et oddetall så får en et oddetall. Et partall har rest 0 når en deler med 2, mens et oddetall har rest 1, så derfor ser addisjonstabellen modulo 2 slik ut:

$+$	0	1
0	0	1
1	1	0

Tilsvarende ser multiplikasjonstabellen slik ut:

$\cdot$	0	1
0	0	0
1	0	1

3.2 Eksempel (Regning modulo 3)

Addisjonstabellen modulo 3 ser slik ut:

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Tilsvarende ser multiplikasjonstabellen slik ut:

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Kongruensregning kan brukes til å forklare noen kjente regneregler.

3.3 Eksempel

Husk at tverrsummen til et tall er summen av sifrene - tverrsummen til 734 er altså 7+3+4=14. En gammel regel sier at et tall er delelig med 3 hvis og bare hvis tverrsummen til tallet er delelig med 3.

Før vi forklarer hvordan en kan bevise dette, viser vi ideen til beviset i et eksempel. La oss undersøke om 261 er delelig med 3. selvfølgelig kunne vi ha delt med 3 og sett om en får 0 i rest. I stedet skriver vi først tallet som en sum av en ener, seks tiere og to hundre:

$$\begin{aligned} 261 &= 2 \cdot 100 + 6 \cdot 10 + 1 \cdot 1 \\ &= 2 \cdot 10^2 + 6 \cdot 10 + 1 \end{aligned}$$

Regner vi modulo 3 er $10 \equiv 1$, dermed får vi

$$\begin{aligned} 261 &= 2 \cdot 10^2 + 6 \cdot 10 + 1 \\ &\equiv 2 \cdot 1^2 + 6 \cdot 1 + 1 \\ &= 9 \equiv 0, \end{aligned}$$

så 261 har rest 0 modulo 3 og er altså delelig med 3.
Generelt, la oss anta at tallet er $a_n a_{n-1} a_{n-2} \cdots a_1 a_0$, der a_i'ene angir sifrene. Størrelsen til tallet er dermed

$$a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + \cdots + a_1 \cdot 10 + a_0.$$

Modulo 3 er $10 \equiv 1$, så vi får

$$\begin{aligned} & a_n \cdot 10^n + a_{n-1} 10^{n-1} + \cdots + a_1 \cdot 10 + a_0 \\ & \equiv a_n \cdot 1^n + a_{n-1} \cdot 1^{n-1} + \cdots + a_i \cdot 1 + a_0 \\ & \equiv a_n + a_{n-1} + \cdots + a_1 + a_0 \quad \text{modulo } 3, \end{aligned}$$

som viser at tallet og tverrsummen har samme rest når de deles med 3.

3.4 Eksempel (Regning modulo 7)

Multiplikasjon med 5 modulo 7 blir:

·	0	1	2	3	4	5	6
5	0	5	3	1	6	4	2

for eksempel blir $3 \cdot 5 \equiv 1$ modulo 7 fordi $15 : 7$ gir 2 med rest 1.
Hele multiplikasjonstabellen modulo 7 (unntatt multiplikasjon med 0 som selvsagt alltid er 0) blir

·	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

Et par kommentarer er på sin plass til denne tabellen. Akkurat slik som faktorenes orden er likegyldig for vanlig multiplikasjon, er den selvsagt det også i kongruensregning. Derfor er tabellen symmetrisk om diagonalen nedover mot høyre. Tabellen er også symmetrisk om den andre diagonalen, uten at vi skal gå inn på noen forklaring av dette her. Viktigere for oss er at hver vannrett rekke i tabellen er en permutasjon av tallene $1, \ldots, 6$. Det vil si at å multiplisere med et fast tall modulo 7 kan brukes som en krypteringsnøkkel i eksemplene fra kapittel 1. Hvis vi ser nøyere etter finner vi den omvendte permutasjonen til en av linjene i tabellen i en annen linje i tabellen. For eksempel gir muliplikasjon med 5 modulo 7 permutasjonen:

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow, \\ 5 & 3 & 1 & 6 & 4 & 2 \end{array}$$

mens multiplikasjon med 3 gir den omvendte permutasjonen

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow. \\ 3 & 6 & 2 & 5 & 1 & 4 \end{array}$$

Det er ikke så vanskelig å forklare hvorfor: Hvis jeg starter med et tall x og multipliserer det først med 5 og deretter med 3 modulo 7 får jeg

$$x \cdot 5 \cdot 3 = x \cdot (5 \cdot 3) \equiv x \cdot 1 = x \quad \text{modulo } 7.$$

Det vil si at hvis jeg først krypterer ved å multiplisere med 5, så kan jeg dekryptere med å multiplisere med 3 modulo 7. Denne egenskapen til multiplikasjontabellen henger nøye sammen med at vi regner modulo et primtall.

3.5 Eksempel (Regning modulo 6)

Hvis vi multipliserer med 3 modulo 6, får vi 0 dersom vi startet med 0, 2 eller 4, og 3 hvis vi starter med 1, 3, eller 5. Det betyr at multiplikasjon med 3 ikke gir permutasjoner av tallene $1, 2, 3, 4, 5$ som ville tilsvare permutasjonene vi fikk i forrige eksempel. Spesielt er ingen produkter med 3 lik 1 modulo 6. Dette skyldes at 3 går opp i 6. Multiplikasjon med 1 eller 5 modulo 6 ville gi permutasjoner av tallene $1, 2, 3, 4, 5$, men disse er for enkle for vår bruk, så de er ikke så interessante.

Av disse to eksemplene formulerer vi en regel:

Hvis s og t er naturlige tall som har en felles faktor, så fins det ikke noe tall u slik at $s \cdot u \equiv 1$ modulo t.

Omvendt gjelder:

Hvis s og t er naturlige tall som ikke har felles faktor, så fins det et tall u slik at $s \cdot u \equiv 1$ modulo t.

3.6 Potenser.

Hvis vi tar potenser av 3 modulo 7 får vi:

x	0	1	2	3	4	5	6
3^x	1	3	2	6	4	5	1

Legg merke til to ting. For det første at de seks siste potensene danner en permutasjon av tallene $1, ..., 6$. Vi skal se i neste kapittel hvordan denne permutasjonen kan brukes som krypteringsnøkkel.
Legg for det andre merke til at 1 forekommer som en potens av 3 modulo 7. Dersom vi regnet modulo 6, ville ingen potens av 3 være lik 1. Dette er fordi 3 går opp i 6, eller er en faktor i 6. Generelt gjelder det at hvis s og t har en felles faktor så vil ingen potens av s være lik 1 modulo t. Dette følger av regneregelen over. Omvendt gjelder at hvis s og t er naturlige tall som ikke har felles faktor, så vil en potens av s være lik 1 modulo t.
Det ville være fint å vite hvilke potenser s som er lik 1 modulo t. Euler viste en fin setning om slike potenser.
Han definerte først et tall $\varphi(t)$ til hvert naturlig tall t, til å være **antall naturlige tall**

mindre enn t som ikke har felles faktor med t.
La oss finne noen verdier $\varphi(t)$. Hvis p er et primtall, så er det ingen tall mindre enn p som har felles faktor med t, så $\varphi(p) = p - 1$. Dersom t er produkt av to primtall p og q som begge er større enn 2, så kan en vise at $\varphi(t) = \varphi(p \cdot q) = (p-1) \cdot (q-1)$. For eksempel er det lett å sjekke at $\varphi(15) = \varphi(3 \cdot 5) = 2 \cdot 4 = 8$, ved å bruke definisjonen direkte, prøv!
Euler viste setningen:
Hvis s og t er naturlige tall uten felles faktor så er $s^{\varphi(t)} \equiv 1$ modulo t.
Denne skal vi få bruk for i neste kapittel.

Oppgaver

1. Lag en addisjons- og en multiplikasjonstabell for for regning modulo 5.

2. Lag en multiplikasjonstabell for regning modulo 6.

3. Lag en multiplikasjonstabell for regning modulo 11.

4. Vis at 9 deler et tall hvis og bare hvis det deler tverrsummen.

5. Den *alternerende tverrsummen* til et naturlig tall n er definert som

$$a_0 - a_1 + a_2 - a_3 + \cdots + (-1)^k a_k$$

   der $n = a_k a_{k-1} a_{k-2} \cdots a_1 a_0$ er tallet skrevet i titallsystemet. Vis at 11 deler n hvis og bare hvis 11 deler den alternerende tverrsummen.

6. Undersøk om 778431276659113 er delelig med 3, 9 eller 11.

4. Hemmelige koder med kongruensregning

Kongruensregningen i forrige kapittel gir oss krypteringsnøkler som er enkle å bruke. De krypteringsnøklene som vi bruker er alle permutasjoner, og slike fant vi igjen nettopp i kongruensregningen.

4.1 Eksempel

Hvis alfabetet vårt består av seks tegn og vi har kodet disse med tallene $1, \ldots, 6$ så kan vi bruke multiplikasjon modulo 7 til å lage krypteringsnøkler. For eksempel gir muliplikasjon med 5 modulo 7 krypteringsnøkkelen:

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 5 & 3 & 1 & 6 & 4 & 2 \end{array},$$

mens multiplikasjon med 3 gir den omvendte permutasjonen, altså dekrypteringsnøkkelen:

x	1	2	3	4	5	6
$3x$	3	6	2	5	1	4

Anta nå at vi fikk en melding som vi visste var kryptert med en slik krypteringsnøkkel, det vil si at vi visste at den var kryptert med multiplikasjon modulo 7, men at vi ikke visste hvilket tall som en hadde multiplisert med. Siden dekrypteringsnøkkelen er multiplikasjon med et annet tall modulo 7, er det bare å prøve et tall om gangen og se om meldingen gir mening. Multiplikasjon modulo 7 er lettere enn å huske en lang permutasjon, så derfor tar det ikke så lang tid å finne den riktige dekrypteringsnøkkelen. Derfor er heller ikke disse krypteringsnøklene særlig sikre.

4.2 Eksempel

Litt sikrere krypteringsnøkler får en ved å ta potenser. I vårt eksempel tar vi potenser modulo 7. Hvis vi tar potenser av 3 modulo 7 får vi:

x	0	1	2	3	4	5	6
3^x	1	3	2	6	4	5	1

Legg merke til at de seks siste potensene danner en permutasjon modulo 7. Hvis en bruker denne som krypteringsnøkkel, så er der ikke noen tilsvarende måte å lage dekrypteringsnøkkelen på. En må rett og slett skrive opp den omvendte permutasjonen fra den

opprinnelige. I eksempelet er antall sifre så lite at dette går greitt. Men når antall sifre blir stort tar dette tid, og gjør derfor disse krypteringsnøklene sikrere enn de en får fra multiplikasjon.

De to første eksemplene viser krypteringsnøkler som baserer seg på at avsender og mottaker begge to kjenner krypteringsnøkkelen og dekrypteringsnøkkelen. I det neste eksempelet, som er mer avansert, er det bare mottaker som kjenner dekrypteringsnøklene, mens avsender (og faktisk alle andre) kjenner krypteringsnøklene. Dette er et mer avansert eksempel, som er i praktisk bruk.

4.3 RSA-systemet.

Kongruensregning inkludert Eulers setning blir brukt i et kryptografisk system som er kjent under betegnelsen RSA (oppkalt etter Rivest, Shamir og Adleman som beskrev metoden i 1977). Vi skal se på en enkel variant av dette systemet. Elevene i en klasse vil sende hemmelige meldinger til hverandre. For enkelhets skyld antar vi at meldingene som skal sendes er tall (det kan også være en tekst som er kodet til et tall etter en felles kodenøkkel). Systemet er slik at alle kan sende hemmelige meldinger til hverandre, som likevel er slik at bare mottakeren kan dekryptere meldingen.

Hver elev velger seg ut to primtall p og q som begge er større enn 2 (i praksis mye større). La $n = pq$. Når p og q er veldig store er det svært vanskelig og i praksis umulig å finne faktoriseringen $n = pq$ når faktorene ikke er kjent. Hver elev regner ut $\varphi(n) = (p-1)(q-1)$ og velger to tall a, b slik at $a \cdot b \equiv 1$ modulo $\varphi(n)$. Da er altså $a \cdot b = k \cdot \varphi(n) + 1$ der k er et naturlig tall.
Tallene n og b gjøres kjent for alle, det skrives for eksempelet på tavla. Når man ikke kjenner p og q er det umulig å beregne $\varphi(n)$ og derfor umulig å finne a selv om b er kjent. Krypteringsnøkkelen er nå å ta b-te potens modulo n. Altså

$$x \mapsto x^b \quad \text{modulo } n$$

mens dekrypteringsnøkkelen er gitt ved å ta a-te potens modulo n, altså

$$y \mapsto y^a \quad \text{modulo } n.$$

Legg merke til at alle elevene kan kryptere (fordi n og b er kjent for alle), mens det bare er den rette mottakeren som kjenner a og som derfor kan dekryptere en hemmelig melding.
Grunnen til at dekrypteringsnøkkelen virkelig dekrypterer ligger i Eulers setning. Hvis x er den opprinnelige meldingen blir den kryperte meldingen x^b modulo n, mens den dekrypterte meldingen blir

$$(x^b)^a = x^{a \cdot b} = x^{k \cdot \varphi(n)+1} = (x^{\varphi(n)})^k \cdot x^1 \equiv x.$$

I den siste overgangen har vi brukt Eulers teorem til å beregne

$$x^{\varphi(n)} \equiv 1$$

Vi skal konkretisere hele prosedyren i et eksempel.

4.4 Eksempel

Vi lar Per velge de odde primtall $p = 7$ og $q = 13$. Det gir $n = 91$ og $\varphi(91) = 6 \cdot 12 = 72$. I tillegg velger han $a = 5$ og $b = 29$ fordi han har regnet ut at $5 \cdot 29 = 145 = 72 \cdot 2 + 1 \equiv 1$ modulo 72. Per gjør tallene $(n, b) = (91, 29)$ kjent for alle elevene i klassen. Hvis Kari nå ønsker å sende meldingen $x = 11$ til Per, regner hun ut 11^{29} modulo 91. Dette kan se uoverkommelig ut, men det fins en snarvei, nemlig å bruke gjentatte kvadreringer. Da skriver hun først eksponenten som en sum av potenser av 2 og utfører deretter kvadreringer modulo 91. Hun får

$$\begin{aligned} 11^{29} &= 11^{16+8+4+1} \\ &= 11^{2^4} \cdot 11^{2^3} \cdot 11^{2^2} \cdot 11^{2^0} \\ &= (((11^2)^2)^2)^2 \cdot ((11^2)^2)^2 \cdot (11^2)^2 \cdot 11 \\ &\equiv 81 \cdot 9 \cdot 81 \cdot 11 \\ &\equiv 72 \quad \text{modulo } 91 \end{aligned}$$

Denne meldingen er den hun sender. En tredje elev som snapper opp den hemmelige meldingen kjenner nå de tre tallene $91, 29, 72$. Det er svært komplisert, i praksis umulig (når tallene er store nok) å finne de tre tallene p, q, a og den opprinnelige meldingen x dersom en bare kjenner disse tre tallene. Per derimot, kjenner p, q, a, men ikke x. Den kan han imidlertid finne ved å regne ut

$$72^a = 72^5 \equiv 11 \quad \text{modulo } 91$$

gjerne ved å bruke kvadreringsoppskriften over og oppsplittingen

$$5 = 2^2 + 2^0.$$

Dermed har Per dekryptert den hemmelige meldingen.

5. Feilrettingskoder og kongruensregning.

Vi skal se på tre eksempler på feilrettingskoder der en bruker kongruensregning til å oppdage og eventuelt rette feil i mottatte meldinger eller **kodeord**. I disse eksemplene regner en modulo 2, 3 eller 11.

5.1 Eksempel

I dette eksempelet består et **kodeord** av fem sifre som alle er 0 eller 1. Det vil si at kodeordet er representert som et 5-sifret tall $x_1x_2x_3x_4x_5$ i 2-tallsystemet. Et slikt tall er et kodeord dersom

$$x_1 + x_3 + x_4 \equiv 0 \quad \text{modulo } 2$$

og

$$x_1 + x_2 + x_3 + x_5 \equiv 0 \quad \text{modulo } 2$$

Her kan en skrive opp alle kodeordene. Et eksempel på et kodeord er 01110. Prøv å skrive opp alle de andre! Legg merke til at to forkjellige kodeord er forskjellige på minst to plasser. Det betyr at dersom det er oppstått en feil, det vil si at et siffer er feil i et kodet melding, så kan en **oppdage** at der er oppstått en feil. Men en kan ikke rette feilen. La for eksempel

$$11011$$

være en mottatt melding. Dette er ikke et kodeord, mens både

$$11010 \quad \text{og} \quad 10011$$

er det. Forskjellen mellom kodeordene og 11011 er bare i ett siffer. Hvis vi antar at det bare er oppstått en feil i den mottatte meldingen, så kan begge disse to ha vært den opprinnelige meldingen. Vi kan imidlertid ikke vite hvilket, så vi kan ikke **rette** feilen selv om vi kunne **oppdage** den.

5.2 Eksempel

I dette eksempelet består et **kodeord** av fire sifre som alle er 0, 1 eller 2. Det vil si at kodeordet er representert som et 4-sifret tall $x_1x_2x_3x_4$ i 3-tallsystemet. Et slikt tall er et kodeord dersom

$$2x_1 + x_2 + x_3 \equiv 0 \quad \text{modulo } 3$$

og

$$2x_1 + 2x_2 + x_4 \equiv 0 \quad \text{modulo } 3$$

Her kan en skrive opp alle kodeordene. Et eksempel er 1220. Prøv å skrive opp alle de andre! To forkjellige kodeord i dette eksempelet er forskjellige på minst tre plasser. Det betyr at dersom det er oppstått en feil, det vil si at et siffer er feil i en kodet melding, så kan en både **oppdage** at det er oppstått en feil og **rette** den.

5.3 ISBN-koden.

Alle bøker klassifiseres ved hjelp av den så kalte ISBN-koden (International Standard Book Number). Denne koden består av et 10-sifret tall representert i 11-tallsystemet, det vil si 10 sifre som hver er en restklasse modulo 11. La oss kalle sifrene $x_1x_2\ldots x_{10}$. Da er sifrene i en ISBN-kode alltid slik at

$$x_1 + 2x_2 + 3x_3 + \ldots + 10x_{10} \equiv 0 \quad \text{modulo } 11.$$

Denne ligningen kalles kontrolligningen, og brukes til å kontrollere om et 10-sifret tall i 11-tallsystemet er et ISBN-tall. Hvis et siffer blir forandret i en ISBN-kode vil kontrolligningen oppdage denne. Hvis vi i tillegg vet i hvilken posisjon feilen er oppstått, kan en bruke kontrolligningen til å korrigere feilen, prøv! Videre vil kontrolligningen også oppdage om to sifre er byttet om, prøv! ISBN-koden er derfor et eksempel på en feilrettingskode.

5.4 Norske personnummer.

I Norge har vi et personnummersystem som i ulik sammenheng blir brukt til å identifisere personer. Når vi blir født, får vi tildelt et 11-sifret tall. De seks første sifrene er fødselsdatoen. De neste tre er vårt personlige nummer, som skal skille mennesker født på samme dato. De siste to sifrene er kontrollsifre. Dersom de første ni sifrene er $x_1x_2\ldots x_9$, så er de tiende sifferet

$$x_{10} \equiv 8x_1 + 4x_2 + 5x_3 + 10x_4 + 3x_5 + 2x_6 + 7x_7 + 6x_8 + 9x_9 \quad \text{modulo } 11$$

mens det ellevte sifferet er gitt ved

$$x_{11} \equiv 6x_1 + 7x_2 + 8x_3 + 9x_4 + 4x_5 + 5x_6 + 6x_7 + 7x_8 + 8x_9 + 9x_{10} \quad \text{modulo } 11$$

De to siste sifrene regner vi altså ut modulo 11. De ni første sifrene er selvsagt enkeltsifre. Men når en regner modulo 11, kan vi få hva som helst mellom 0 og 10. De to kontrollsifrene vil dermed også kunne være 2-sifrede (10). For å unngå dette passer man på å velge de tre sifrene i det personlige nummeret slik at de hverken x_{10} eller x_{11} blir 10.

Oppgaver

1. Skriv opp alle kodeordene i eksempel 1.

2. Vis at alle enkeltfeil i et norsk personnummer kan oppdages og korrigeres selv om vi ikke vet hvor feilen er oppstått.

3. Vis at alle feil som skyldes ombytting av to siffer i et norsk personnummer kan oppdages.

6. Undervisningsopplegg.

6.1 HEMMELIGE KODER I KLASSEROM

Disposisjon for 1-2 timer:

1. Introduksjon med kryptogrammer

2. Kode- og krypteringsnøkler (alfabet og permutasjoner)

3. Oppsett for bruk av (symmetrisk) kryptering i kommunikasjon

4. Effektiv krypering med kongruensregning

5. Regneøvelser med kongruensregning

6. Regneøvelser med krypterings og dekrypteringsnøkler med kongruensregning.

7. Oppgaver der en med gitt alfabet og kryptert melding skal finne meldingen(når krypteringsnøkkelen er gitt ved multiplikasjon med et tall modulo t som i punkt 4)

8. Oppsummering om inverser i moduloregning

9. Øvelse med å kryptere og dekryptere en melding

OPPGAVER UNDERVEIS:

1. Lag en regnetabell for muliplikasjon modulo 7

2. Lag dekryperingsnøkler til krypteringsnøklene i oppgave 1. Finner du noe mønster?

3. Gå sammen to og to. Lag hver for seg en melding med 6 forskjellige bokstaver. Skriv bokstavene i alfabetisk rekkefølge. Velg en hemmelig krypteringsnøkkel med multiplikasjon modulo 7. Krypter den opprinnelige meldingen med denne nøkkelen. Send den krypterte meldingen, sammen med bokstavene du har brukt i alfabetisk rekkefølge til hverandre. Forsøk å finne ut hvilken krypteringsnøkkel som den andre brukte og dekrypter meldingen.

6.2 FEILRETTINGSKODER I KLASSEROM

Disposisjon for 1-2 timer:

1. Introduksjon om kodeord og feil i kodeord.

2. Koding av meldinger som oppdager og retter feil (repetisjonskoder).

3. Kongruensregning modulo 2 og 3.

4. Effektive feilrettingskoder med kongruensregning (eksempel 5.1 og 5.2)

5. Oppgaver der en med gitt(e) kontrolligning(er) skal finne feilen.

6. Det norske personnummersystemet.

OPPGAVER UNDERVEIS:

1. Finn alle kodeordene for en feilrettingskode laget med sifre modulo 2 (eksempel 5.1).

2. Finn alle kodeordene for en feilrettingskode laget med sifre modulo 3 (eksempel 5.2).

3. Lag et personnummer etter det norske personnummersystemet.

4. Gå sammen to og to. Lag hver for seg en melding med en feil i en feilrettingskode og prøv og finn feilen hos den andre (bruk gjerne kodene i eksempel 5.1 eller 5.2).

7. Prosjektoppgaver.

7.1 RSA-systemet for utveksling av hemmelige koder

Område: Aritmetikk

Klassetrinn: Grunnkurs 1MY

1. Gjør rede for hvordan RSA systemet kan brukes til å utveksle hemmelige meldinger mellom et stort antall personer.

2. Lag en modell av RSA systemet der en bare bruker relativt små primtall (mindre enn 20) for bruk mellom elevene i en klasse.

I redegjørelsen kan blant annet inngå nedenstående punkter:
- Potenser modulo et (lite) naturlig tall (gjerne eksempler).
- Eulers φ-funksjon, hva er denne funksjonen, eksempler.
- Krypteringsnøkler og dekrypteringsnøkler i RSA systemet.
- Eksempler på hvordan RSA systemet kan brukes i en klasse.

Du kan også løse en av følgende oppgaver:

Oppgave 1: Undersøk Eulers φ-funksjon. Hva er $\varphi(7 \cdot 11), \varphi(3 \cdot 7 \cdot 11)$? Hvis p, q, r er tre primtall, finn en formel for $\varphi(p \cdot q \cdot r)$?

Oppgave 2: Regn ut $7^6 3, 13^{111}, 17^{257}$ modulo 11.

Kilder:
B. Johnsen: Kryptografi - en gammel disiplin med moderne anvendelser, i Per Hag & Ben Johnsen (red.): *Fra Matematikkens Spennende Verden*, Tapir, 1993, 123-134.

7.2 Hemmelige koder og andre verdenskrig

Område: Aritmetikk og historie

Klassetrinn: Grunnkurs 1MY

Gjør rede hemmelige koder som ble brukt av det norske militæret under de andre verdenskrig.

I redegjørelsen kan blant annet inngå nedenstående punkter:
- Beskrivelse av det utstyret som ble brukt.
- Eksempler på kodenøkler.
- Bokkoden. Eksempel med lange kodenøkler hentet fra en valgt bok.

Du kan også løse følgende oppgave:

Oppgave 1: Dette er en melding som er kodet etter bokkoden med første linje i nasjonalsangen som kodenøkkel.

PØMHACSAGOAEHTNGTTJPGNU

Finn den opprinnelige meldingen.

Kilder:
B. Johnsen: Kryptografi - en gammel disiplin med moderne anvendelser, i Per Hag & Ben Johnsen (red.): *Fra Matematikkens Spennende Verden*, Tapir, 1993, 123-134.
Forsvarsmuseet, Hjemmefrontmuseet.

7.3 Enigma

Område: Aritmetikk og historie

Klassetrinn: Grunnkurs 1MY

Beskriv den tyske krypteringsmaskinen, hvordan den fungerer og litt av historien rundt.

I redegjørelsen kan blant annet inngå nedenstående punkter:
- Beskrivelse av utstyret.
- Beskrivelse av hvordan meldinger blir kryptert med maskinen.
- Eksempler med en forenklet modell (se Singhs bok) på kryptering.

Du kan også løse følgende oppgave:

Oppgave 1: Vis ved regning hvor mange krypteringsnøkler (innstillinger) der er på en Enigma maskin. (Her må du velge modell, eventuelt gi antallet for hver modell).

Kilder:
S. Singh: Koder, Aschehoug 2000
Forsvarsmuseet

7.4 Det norske personnummersystemet

Område: Aritmetikk

Klassetrinn: Grunnkurs 1MY

Beskriv det norske personnummersystemet. Prøv deretter å lage et eget 6 sifret system der fødselsdag (ukedag) og kjønn er gitt, et personnummer velges (ett eller to sifre) og ett (eller to) kontrollsiffer lages der en regner modulo 2, 3.

I redegjørelsen kan blant annet inngå nedenstående punkter:
- Regning modulo 11, multiplikasjonstabell.
- Forklar hvordan de fem siste sifrene i personnummeret blir laget.
- Forklar hvordan feil i ett siffer et personnummer kan oppdages og korrigeres.

Lag et personnummer basert på ukedag og kjønn i 3-tall systemet. De to første sifrene bestemmer ukedag, det neste bestemmer kjønn. De tre siste sifrene, er to fritt valgte og ett kontroll siffer. Prøv å finne en ligning for det kontrollsifferet slik at koden for oppdager en feil.

Du kan også løse følgende oppgave:

Oppgave 1: **01010101013** er et personnummer med en feil. Finn og korriger denne feilen.

Oppgave 2: Forklar hvordan ombytting av to siffer i et personnumer kan oppdages.

Kilder:
Kapittel 5 i dette heftet.

7.5 Strekkoden

Område: Aritmetikk

Klassetrinn: Grunnkurs 1MY

Beskriv hvordan strekkoden som alle varer i buutikker i er merket med, er laget. Lag deretter en egen strekkode med fem streker og to tykkelser som oppdager feil.

I redegjørelsen kan blant annet inngå nedenstående punkter:
- Finn ut hvilke tykkelser og mellomrom som er brukt i strekkoden.
- Er der sammenheng mellom strekene og tallene som står ved siden av?
- Forklar hvordan feil strekkode oppdages.
- Finn ut hvem som lager strekkodene.

Du kan også løse følgende oppgave:

Oppgave 1: Skriv opp alle kodeordene i eksempel 5.1 og 5.2 i heftet.

Oppgave 2: Vis at koden i eksempel 5.2 oppdager og korrigerer en feil

Kilder:
Kapittel 5 i dette heftet.

7.6 ISBN-koden

Område: Aritmetikk

Klassetrinn: Grunnkurs 1MY

Emne: 2.

Gjør rede for hvordan ISBN-koden som alle bøker er merket med fungerer, finn ut hvem som lager den og hvordan den brukes.

I redegjørelsen kan blant annet inngå nedenstående punkter:
- Regning modulo 11.
- Hvordan kan en bruke kontrolligningen til å finne feil i en ISBN-kode.
- Hvem produserer ISBN-koder for bøker, når i bokproduksjonen blir dette nummeret laget?

Du kan også løse følgende oppgave:

Oppgave 1: **0387996375** er en ISBN-kode med feil 6. siffer. Hva er det riktige sifferet?

Oppgave 2: Forklar hvorfor kontrolligningen kan oppdage at to sifre er byttet om i en ISBN-kode.

Kilder:
Kapittel 5 i dette heftet.

I dette hefte har vi holdt oss til anvendelser av enkel kongruensregning i kodeteori og kryptologi. To gode artikler for videre lesning om disse temaene er skrevet av Ben Johnsen [1] og Leif Nilsen [3]. Boka til Singh [4] gir en populær historisk oversikt over bruk av hemmelige koder. Artikkelen til van Lint [2] tar for seg de feilrettingskodene som brukes på CD-er.

Referanser

1. B. Johnsen: Kryptografi - en gammel disiplin med moderne anvendelser, i Per Hag & Ben Johnsen (red.): *Fra Matematikkens Spennende Verden*, Tapir, 1993, 123-134.
2. Jacobus H. van Lint: Kompaktskivans matematik, *Normat*, **48** (2000), 115-122.
3. L. Nilsen: Modulære kvadratrøtter og moderne kryptologi, *Normat*, **40** (1992), 75-89.
4. S. Singh: Koder, Aschehoug 2000

Løsninger på eksempel 5.1 og 5.2

5.1 Eksempel.

I dette eksempelet består et **kodeord** av fem sifre som alle er 0 eller 1. Det vil si at kodeordet er representert som et 5-sifret tall $x_1x_2x_3x_4x_5$ i 2-tallsystemet. Et slikt tall er et kodeord dersom

$$x_1 + x_3 + x_4 \equiv 0 \quad \text{modulo } 2$$

og

$$x_1 + x_2 + x_3 + x_5 \equiv 0 \quad \text{modulo } 2$$

Kodeordene er da:

0	0	0	0	0
0	1	1	1	0
1	0	1	0	0
1	1	0	1	0
1	0	0	1	1
0	1	0	0	1
1	1	1	0	1
0	1	1	1	1

5.2 Eksempel.

I dette eksempelet består et **kodeord** av fire sifre som alle er 0, 1 eller 2. Det vil si at kodeordet er representert som et 4-sifret tall $x_1x_2x_3x_4$ i 3-tallsystemet. Et slikt tall er et kodeord dersom

$$2x_1 + x_2 + x_3 \equiv 0 \quad \text{modulo } 3$$

og

$$2x_1 + 2x_2 + x_4 \equiv 0 \quad \text{modulo } 3$$

Kodeordene er da:

0	0	0	0
0	1	2	1
0	2	1	2
1	0	1	1
1	1	0	2
1	2	2	0
2	0	2	2
2	1	1	0
2	2	0	1