

De tweede ciphertekst was versleuteld met Playfair. We moesten dus de key zien te vinden waarmee een matrix was opgesteld om digrammen te versleutelen. We hadden op dit punt al een Nederlandse en een Franse tekst ontcijferd, dus we gingen ervanuit dat dit Engels was. Onze eerste poging bestond uit een frequentieanalyse van de digrammen. We vergeleken die waarden met gekende waarden voor digrammen in het Engels die we zelf hadden berekend aan de hand van enkele Engelse plainteksten. Na veel puzzelwerk raakten we echter niet erg ver. Er zijn 600 digrammen in Playfair (hoewel ze niet allemaal in de ciphertekst voorkwamen) en door dicht bij elkaar gelegen frequenties was het bijna onmogelijk ze juist te kiezen. We bekeken ook de frequenties van opeenvolgende digrammen samen, we hielden rekening met frequente digrammen waarvan ook het omgekeerde frequent was en probeerden de ciphertekst te bruteforcen met keys waarbij enkele gevonden digrammen klopten, maar vonden ook zo het antwoord niet. Daarna gooiden we het over een andere boeg en gingen we op zoek naar een algoritme om de ciphertekst volledig geautomatiseerd te kraken. Eerst maakten we een implementatie van een hill climb algoritme, maar ook hiermee vonden we geen oplossing. Daarna stootten we op het zogenaamde Churn algoritme^{1 2}. Dit algoritme is gelijkaardig aan simulated annealing, alleen heel wat simpeler om te implementeren.

Elke plaintekst krijgt een score toegewezen als volgt: voor elke mogelijke digram in het Engels werd de frequentie geanalyseerd. Hiervan werden telkens de log genomen om de invloed van erg grote waarden te beperken, en werden deze waarden herschaald naar 0-9. Nu wordt voor elk digram (ook twee opeenvolgende letters die volgens Playfair in verschillende digrams zitten!) in de plaintekst de frequentie van 0-9 bij de score van de plaintekst geteld. Hoe hoger de score, hoe waarschijnlijker dat de tekst Engels is dus. Merk hierbij op dat de plainteksten van de gegeven ciphertekst komen, dus erg hoog scorende maar niet-Engelse plainteksten als "eeeeeeeeee" zullen niet voorkomen. In het algoritme starten we met een zogenaamde parent key (bv. gewoon het alfabet) waarmee we de ciphertekst ontcijferen en een score toekennen. Daarna voeren we een kleine verandering door aan de parent key en noemen we het resultaat de child key. Deze verandering kan een horizontale of verticale spiegeling, een combinatie van de twee, een verwisseling van twee willekeurige rijen of kolommen of een verwisseling van twee letters zijn. Hierbij komt de verwisseling van twee letters veel meer voor dan de andere opties, dit omdat er veel meer mogelijke verwisselingen van twee letters zijn. De plaintekst voor de child key wordt ook geëvalueerd. Indien de child key beter scoorde, vervangt deze de parent key. Indien de parent key beter scoorde, wordt een willekeurig getal uit een array van 100 getallen gekozen.

¹<http://www.cryptoden.com/index.php/algorithms/churn-algorithm/20-churn-algorithm>

²<http://s13.zetaboards.com/Crypto/topic/6781204/1/>

Indien het verschil tussen parent en child key scores minder was dan dit gekozen getal, vervangt de child key toch de parent key. Hierdoor kan het algoritme uit lokale maxima raken. Het algoritme blijft oneindig lopen en print de uitkomst van een iteratie enkel indien een nieuwe topscore bereikt is. Merk op dat de 100 getallen in de genoemde array zodanig gekozen zijn dat de kans dat child parent vervangt, gelijkaardig is aan die bij simulated annealing. Bij sommige runs van het algoritme vonden we al na een tweeduizendtal iteraties een tekst die heel erg op Engels leek. Het antwoord was niet helemaal correct gezien bij de logaritmen van de frequenties van digrammen geen rekening gehouden werd met de meer voorkomende X bij Playfair. Het was wel dicht genoeg bij Engels dat we de laatste aanpassingen handmatig konden doorvoeren. We vonden als key "A brief history of time", met als plaintekst het begin van "A Brief History of Time: From the Big Bang to Black Holes" door Stephen Hawking.³ Merk op dat we de twee lijsten met waarden op <http://www.cryptoden.com/> vonden, maar de code verder helemaal zelf geschreven is en enkel gebaseerd is op de beschrijving van het algoritme.

Volgens de beschrijving van het algoritme waren de 100 waarden in de array afgesteld op een ciphertekst van 100-120 karakters en moest deze nog geschaald worden voor langere teksten. Wij kregen echter betere resultaten zonder de waarden te schalen.

Het algoritme in playfair.py is een licht aangepaste versie van wat we eerst gebruikten. Het geeft vaker snel een antwoord door rollbacks uit te voeren als een tijdje geen vooruitgang geboekt wordt, of zelfs helemaal opnieuw te beginnen. Ook is het scoren nu aangepast aan de X'en in Playfair waardoor exact de correcte plaintekst wordt gevonden. De gevonden key matrix is niet altijd de matrix gegenereerd met "abriefhistoryoftime", maar wel altijd een correcte matrix. Gezien niet alle digrammen voorkomen, zijn meerdere correcte matrices mogelijk.

In de eerste twee lijntjes van het Churn algoritme wordt een vaste seed ingesteld. Deze levert al na 2406 iteraties het antwoord op. Om met een andere seed te proberen, moeten de eerste twee regels code verwijderd worden. Hierdoor kan het algoritme wel langer duren. De variërende runtime is een gevolg van de niet-deterministische aard van het algoritme. Gezien de code niet concurrent is, is het aangewezen om het programma meerdere keren tegelijk te draaien om sneller tot een resultaat te komen.

AWELXLKNOWNSCIENTISTSOMESAYITWASBERTRANDRUSX
SELXLONCEGAVEAPUBLICLECTUREONASTRONOMYHEDESCRIB
EDHOWTHEXEARTHORBITSAROUNDTHESUNANDHOWTHESUNI
NTURNORBITSAROUNDTHECENTEROFAVASTCOLXLECTIONOF
STARSCALXLEDOURGALAXYATXTHEXENDOFTHELECTUREALI

³http://www.fisica.net/relatividade/stephen_hawking_a_brief_history_of_time.pdf#page=3

TXTLEOLDLADYATXTHEBACKOFTHEROXOMGOTUPANDSAIDW
 HATYOUHAVETOLDUSISRUBXBISHTHEWORLDISREALXLYAFLA
 TPLATESUPXPORTEDONTHEBACKOFAGIANTXTORTOISETHES
 CIENTISTGAVEASUPERIORSMILEBEFOREREPLYINGWHATISTH
 ETORTOISESTANDINGONYOUREVERYCLEVERYOUNGMANVER
 YCLEVERSAIDTHEOLDLADYBUTITSTURTLESALXLTHEWAYDO
 WNMOSTPEOPLEWOULDFINDTHEPICTUREOF FOURUNIVERSEAS
 ANINFINITETOWEROFTORTOISESRATHERXRIDICULOUSBUTW
 HYDOWETHINKWEKNOWBETXTERWHATDOWEKNOWABOUTX
 THEUNIVERSEANDHOWDOWEKNOWITWHERE DIDTHEUNIVERS
 ECOMEFROMANDWHEREISITGOINGDIDTHEUNIVERSEHAVEABE
 GINXNINGANDIFSOWHATHAPXPENEDBEFORETHENWHATISTH
 ENATUREOFTIMEWILXLITEVERCOMETOANENDCANWEGOBAC
 KINTIMERECENTBREAKTHROUGHSINPHYSICSMADEPOSXSIBLE
 INPARTBYFANTASTICNEWTECHNOLOGIESXSUGXGESTANSWE
 RSTOSOME OFTHESE LONGSTANDINGQUESTIONSXSOMEDAYTH
 ESEANSWERSMAYSEXEMASOBVIOUSTOUSASTHEXEARTHORBI
 TINGTHESUNORPERHAPSASRIDICULOUSASATOWEROFTORTOI
 SESONLYTIMEWHATEVERTHATMAYBEWILXLTELXLASLONGAG
 OASTHREXEHUNDREDANDFOURTYBCTHEGREXEKPHILOSOPH
 ERARISTOTLEINHISBOXOKONTHEHEAVENSWASABLETOPUTFO
 RWARDTWO GOXOD ARGUMENTS FORBELIEVINGTHATXTHEXE
 ARTHWASAROUNDSPHERERATHERTHANAHATPLATEFIRSTHER
 EALIZEDTHATECLIPSESOFTHEMOXONWERECAUSED BYTHEXE
 ARTHCOMINGBETWEXENTHESUNANDTHEMOXONTHEXEARTH
 SXSHADOWONTHEMOXONWASALWAYSROUNDWHICHWOULDB
 ETRUEONLYIFTHEXEARTHWASXS SPHERICALIFTHEXEARTHXHA
 DBEXENAFLATDISKTHESHADOWXWOULDHAVEBEXENELONGA
 TEDANDELXLIPTICALUNLESXSTHEXECLIPSEALWAYSOCXCURX
 REDATATIMEWHENTHESUNWASDIRECTLYUNDERTHECENTER
 OFTHEDISKSECONDTHEGREXEKSKNEWFROMTHEIRTRAVELST
 HATXTTHENORTHSTARAPXPEAREDLOWERINTHESKYWHENVIE
 WEDINTHESOUTHTHANITDIDINMORENORTHERLYREGIONSXS
 NCETHENORTHSTARLIESOVERTHENORTHPOLEITAPXPEARSTO
 BEDIRECTLYABOVEANOBSERVERATXTTHENORTHPOLEBUTXTO
 SOMEONELOXOKINGFROMTHEXEQUATORITAPXPEARSTOLIEIU
 STATXTHEHORIZONFROMTHEDIFXERENCEINTHEAPXPARENT
 POSITIONOFTHENORTHSTARINEGYPTANDGREXECEARISTOTL
 EXEVENQUOTEDANESTIMATE THATXTTHEDISTANCEAROUNDTH
 EXEARTHWASFOURHUNDREDTHOUSANDSTADIAITISNOTKNOW
 NEXACTLYWHATLENGTHASTADIUMWASBUTITMAYHAVEBEXE
 NABOUTXTTWOHUNDREDDYARDSWHICHWOULDMAKEARISTOTL
 ESESTIMATEABOUTXTTWICETHECURXRENTLYACXCEPTEDFIG
 URETHEGREXEKSEVENHADATHIRDARGUMENTXTTHATXTHEXE

ARTH MUST BE ROUND FOR WHY ELSE DOES ONE FIRST SEE THE SAIL
OF A SHIP COMING OVER THE HORIZON AND ONLY LATER SEE THE
HULL? LARISTOTLE THOUGHT THE EARTH WAS STATIONARY AND
THAT THE SUN, THE MOON, THE PLANETS AND THE STARS
MOVED IN CIRCULAR ORBITS ABOUT THE EARTH. HE BELIEVED
THIS BECAUSE HE FELT FOR MYSTICAL REASONS THAT THE EARTH
WAS THE CENTER OF THE UNIVERSE AND THAT CIRCULAR MOTION
WAS THE MOST PERFECT. THIS IDEA WAS ELABORATED BY PTOLEMY
IN THE SECOND CENTURY AD INTO A COMPLETE COSMOLOGICAL
MODEL. THE EARTH STOOD AT THE CENTER, SURROUNDED
BY EIGHT SPHERES THAT CARRIED THE MOON, THE SUN, THE
STARS AND THE FIVE PLANETS KNOWN AT THE TIME: MERCURY, VENUS,
MARS, JUPITER AND SATURN.