Andrzej M. Borzyszkowski
Błędne powtórzenie klucza jednorazowego

Zasadą szyfru jednorazowego jest **niepowtarzalność** klucza. E(k,m)=k⊕m; D(k,c)=k⊕c. Jeśli dwie wiadomości są zaszyfrowane tym samym kluczem k, to można obliczyć m1⊕m2=c1⊕c2. Znajomość xor dwu wiadomości niesie już pewne informacje, a każda uzyskana informacja jest z definicji złamaniem szyfru, nawet jesli nie prowadzi do całkowitego odszyfrowania tekstu.

W zadaniu pokazującym możliwości uzyskania informacji ze znajomości xor założymy, że szyfrowane są wyłącznie litery i spacje, być może dla ułatwienia będzie można założyć, że litery są wyłącznie małe. Założymy też, że cały tekst (angielski) jest kodowany standardowo kodem ascii, tzn spacja ma numer 32, a litery 97-122. W notacji heksagonalnej spacja jest równa 0x00100000 a małe litery 0x011..... W sposób xor dwóch liter zaczyna się od trzech zer, a xor litery i spacji ma na początku 010. Wiedząc, że m1⊕m2 ma pierwsze trzy bity 010 wiemy, że jeden ze znaków jest spacją więc m1⊕m2⊕00100000 jest drugim ze znaków, nie wiadomo którym. Jeśli mamy do dyspozycji m1⊕m2 i m2⊕m3 i np. pierwsza para ma spację a druga nie ma, to wiadomo, że spacją jest m1, i wyliczamy m2 i m3. Jeśli obie mają spacje, to prawdopodobnie jest to m2 i wyliczamy pozostałe znaki. Inny przypadek byłby możliwy, gdyby m3⊕m1=00000000 czyli m1=m3. Wówczas być może m1 i m3 byłyby spacjami, a m2 jakimś znakiem. Jeśli znamy więcej przykładów kryptogramów powstałych z użyciem tego samego klucza, to jest duża szansa, na odtworzenie dokładnych tekstów.

Zadanie

Program o nazwie xor powinien umożliwiać wywołanie z linijki rozkazowej z następującymi opcjami:

-p (przygotowanie tekstu do przykładu działania),
-e (szyfrowanie),
-k (kryptoanaliza wyłącznie w oparciu o kryptogram)

Nazwy plików są następujące:

orig.txt: plik zawierający dowolny tekst,
plain.txt: plik z tekstem zawierającym co najmniej kilkanaście linijek równej długości, np. 64,
key.txt: plik zawierający klucz, który jest ciągiem dowolnych znaków podanej wyżej długości,
crypto.txt: plik z tekstem zaszyfrowanym, każda jego linijka jest operacją ⊕ z kluczem,
decrypt.txt: plik z tekstem odszyfrowanym.

Uwaga: pod uwagę będą brane wyłącznie programy z kryptoanalizą.