Andrzej M. Borzyszkowski
Funkcje skrótu

Dwie najpopularniejsze i najczęściej używane funkcje skrótu to **md5** oraz **sha-1**. Na wielu systemach są one dostępne bez dodatkowych instalacji, na komputerze sigma polecenia brzmią md5sum oraz sha1sum. Pierwsza z tych funkcji zwraca skrót 128-bitowy, druga 160-bitowy. Standard SHA udostępnia wiele dalszych funkcji: sha224sum sha256sum sha384sum sha512sum, dających coraz dłuższe skróty. Skróty zapisywane są w systemie szesnastkowym (bez dodatkowych wyjaśnień, że chodzi o ten system zapisu).

Argumentem funkcji jest nazwa pliku, można używać wyrażeń regularnych. Np.

```
[amb@sigma ~]$ md5sum bookmarks.*
```

produkuje komunikat

```
cbeb720b717f7e25791f6c4ed5523d7d   bookmarks.html
2cc3f5956f0a6ec662b0f15452d435ec   bookmarks.zip
```

i podobnie dla funkcji SHA. Komunikat ten można zapisać do pliku, np. poprzez przekierunkowanie. Pozwala to w przyszłości sprawdzić, czy zaszła zmiana:

```
[amb@sigma ~]$ md5sum -c plik
bookmarks.html: NIEPOWODZENIE
bookmarks.zip: DOBRZE
md5sum: UWAGA: 1 z 2 wyliczonych sum kontrolnych się NIE zgadza
```

Funkcje skrótu mogą również czytać dane wejściowe ze standardowego wejścia, np.

```
[andrzej@sigma]$ cat hash.pdf | sha1sum
2af2bbc4c91bcc13dafbea711f9ffa1afa1bb1d0   -
[andrzej@sigma]$ sha1sum hash.pdf
2af2bbc4c91bcc13dafbea711f9ffa1afa1bb1d0 hash.pdf
```

Zadania:

1. Przygotować plik personal.txt ze swoimi danymi osobowymi. Obliczyć wszystkie funkcje skrótu na tym pliku, wyniki zapisać do pliku hash.txt w kolejności coraz dłuższych skrótów.
2. Przygotować drugą wersje pliku z tymi samymi danymi osobowymi personal_.txt, różniącą się jedynie dodatkowym pustym wierszem na końcu. Obliczyć wartość wszystkich funkcji skrótu dla obu wersji pliku połączonego z plikiem pdf wykładu hash.pdf (tzn. wykonać polecenia:

```
3. cat hash.pdf personal.txt | md5sum >> hash.txt
   cat hash.pdf personal_.txt | md5sum >> hash.txt
```

   itd. dla obu wersji pliku z danymi osobowymi). Następnie sprawdzić liczbę bitów (nie bajtów) różnych w obu wynikach. Należy się spodziewać, że w każdej parze ok. połowa bitów będzie różna. Proszę przesłać oba pliki personal.txt oraz plik diff.txt zawierający sześć par wyników dla każdej z funkcji skrótu i liczbę bitów różniących te wyniki (plus program liczący te bity wraz ze źródłem). Nie przesyłać pliku hash.pdf.

   Przykładowy plik z wynikami: diff.txt.