**COEN 350: Network Security**

**ASSIGNMENT 5**

**Shweta Kharat**

**Date: 08/27/2019**

**LZW Algorithm**

**Code (Java):**

**import** java.text.DecimalFormat;

**import** java.util.HashMap;

**import** java.util.Map;

**import** java.util.Map.Entry;

**public** **class** LZW {

**static** Map<Integer, String> *dictionary* = **new** HashMap<>();

**static** **int** []*encode* = **new** **int**[4096];

**static** **int** *cnt* = 0;

**public** **static** **void** main(String[] args) {

// **TODO** Auto-generated method stub

String str = "ACBBAAC";

String str1 = "AAABAABBBB";

String test1 = "\"I believe in intuition and inspiration. Imagination is more "

+ "important than knowledge. For knowledge is limited, whereas imagination "

+ "embraces the entire world, stimulating progress, giving birth to evolution."

+ " It is, strictly speaking, a real factor in scientific research.\" - "

+ "Albert Einstein";

String test2 = "\"I do not know what I may appear to the world, but to myself "

+ "I seem to have been only a boy playing on the seashore, and "

+ "diverting myself in now and then finding a smoother pebble or "

+ "prettier shell than ordinary, whilst the great ocean of truth "

+ "lay all undiscovered before me.\" - Isaac Newton";

**for**(**int** i=0; i<= 31; i++){

*dictionary*.put(i, "0");

}

**for**(**int** i=32; i<=255; i++){

String s1 = Character.*toString*((**char**)(i));

*dictionary*.put(i, s1);

}

//Note: Uncomment encode function for the string to be tested

*LZWEncoding*(str);

//LZWEncoding(str1);

//LZWEncoding(test1);

//LZWEncoding(test2);

*LZWDecoding*(*encode*, *cnt*);

}

**public** **static** **void** LZWEncoding(String str){

System.***out***.println("Uncompresed text = "+ str);

String word = "";

**int** size = 256;

**int** index = 0;

**for**(**int** i=0; i<str.length(); i++){

System.***out***.println("\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*");

System.***out***.println("Iteration " +i);

**char** x = str.charAt(i);

String temp = word + x;

System.***out***.println("Next char X = "+ x);

**if**(*dictionary*.containsValue(temp)){

System.***out***.println(temp + " in dictionary");

word = word + x;

**for** (Entry<Integer, String> entry : *dictionary*.entrySet()) {

**if** (entry.getValue().equals(temp)) {

index = entry.getKey();

}

}

}

**else**{

System.***out***.println(temp + " not in dictionary");

**for** (Entry<Integer, String> entry : *dictionary*.entrySet()) {

**if** (entry.getValue().equals(word)) {

index = entry.getKey();

}

}

*encode*[*cnt*] = index;

*cnt*++;

String s2 = word + x;

*dictionary*.put(size, s2);

System.***out***.println("New entry of dictionary "+ size+

"\t"+*dictionary*.get(size));

size++;

word = Character.*toString*(x);

}

System.***out***.println("word = "+ word);

}

*encode*[*cnt*] = index;

*cnt*++;

System.***out***.print("\nCompressed output is ");

**for**(**int** i=0; i<*cnt*; i++){

System.***out***.print(*encode*[i] + " ");

}

**double** ratio=0;

**int** bitsPerIndex = 16;

**int** bitsPerChar = 8;

System.***out***.println("\nCnt = "+ *cnt*);

System.***out***.println("StrLen = " + str.length());

ratio = (**double**)(*cnt*) / str.length();

System.***out***.println("\nCompression ratio = "+ ratio);

}

**public** **static** **void** LZWDecoding(**int**[] encode, **int** cnt){

System.***out***.println();

**int** size = 256;

Map<Integer, String> dictionary1 = **new** HashMap<>();

**for**(**int** i=0; i<= 31; i++){

dictionary1.put(i, "0");

}

**for**(**int** i=32; i<=255; i++){

String s1 = Character.*toString*((**char**)(i));

dictionary1.put(i, s1);

}

String element = dictionary1.get(encode[0]);

String word = element;

**for**(**int** i=1; i<cnt; i++){

**int** index1 = encode[i];

//System.out.println("Read index "+ index1);

element = dictionary1.get(index1);

**if**(!dictionary1.containsKey(index1)){

//System.out.println("Add new entry in dictionary ");

element = word + word.charAt(0);

}

//System.out.println("Element "+element);

String s4 = word + element.charAt(0);

dictionary1.put(size, s4);

//System.out.println("New entry to dictionary " + size+" "+ s4);

size++;

word = element;

}

System.***out***.println("\nDecompressed output = ");

**for**(**int** i=0; i<cnt; i++){

System.***out***.print(dictionary1.get(encode[i])+"");

}

System.***out***.println("\nNew dictionary entries: ");

**for**(**int** j=256; j<255+cnt; j++){

System.***out***.println(j + " "+dictionary1.get(j));

}

}

}

**Output:**

Uncompresed text = ACBBAAC

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 0

Next char X = A

A in dictionary

word = A

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 1

Next char X = C

AC not in dictionary

New entry of dictionary 256 AC

word = C

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 2

Next char X = B

CB not in dictionary

New entry of dictionary 257 CB

word = B

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 3

Next char X = B

BB not in dictionary

New entry of dictionary 258 BB

word = B

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 4

Next char X = A

BA not in dictionary

New entry of dictionary 259 BA

word = A

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 5

Next char X = A

AA not in dictionary

New entry of dictionary 260 AA

word = A

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 6

Next char X = C

AC in dictionary

word = AC

Compressed output is 65 67 66 66 65 256

Cnt = 6

StrLen = 7

Compression ratio = 0.8571428571428571

Decompressed output =

ACBBAAC

New dictionary entries:

256 AC

257 CB

258 BB

259 BA

260 AA

Uncompresed text = AAABAABBBB

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 0

Next char X = A

A in dictionary

word = A

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 1

Next char X = A

AA not in dictionary

New entry of dictionary 256 AA

word = A

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 2

Next char X = A

AA in dictionary

word = AA

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 3

Next char X = B

AAB not in dictionary

New entry of dictionary 257 AAB

word = B

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 4

Next char X = A

BA not in dictionary

New entry of dictionary 258 BA

word = A

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 5

Next char X = A

AA in dictionary

word = AA

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 6

Next char X = B

AAB in dictionary

word = AAB

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 7

Next char X = B

AABB not in dictionary

New entry of dictionary 259 AABB

word = B

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 8

Next char X = B

BB not in dictionary

New entry of dictionary 260 BB

word = B

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Iteration 9

Next char X = B

BB in dictionary

word = BB

Compressed output is 65 256 66 257 66 260

Cnt = 6

StrLen = 10

Compression ratio = 0.6

Decompressed output =

AAABAABBBB

New dictionary entries:

256 AA

257 AAB

258 BA

259 AABB

260 BB

Uncompresed text = "I believe in intuition and inspiration. Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution. It is, strictly speaking, a real factor in scientific research." - Albert Einstein

\*\*\*\*\*\*\*\*

Compressed output is 34 73 32 98 101 108 105 101 118 101 32 105 110 266 110 116 117 105 116 105 111 268 97 110 100 269 115 112 105 114 97 274 276 46 32 73 109 97 103 267 286 275 268 105 115 32 109 111 114 265 105 109 112 303 116 278 116 32 116 104 278 32 107 110 111 119 108 101 100 103 101 289 70 303 317 319 321 323 325 266 300 261 109 273 323 44 32 119 104 101 304 97 300 306 293 295 287 268 101 109 98 285 99 101 300 314 265 101 270 284 265 119 303 108 100 341 115 274 109 117 108 296 110 103 32 112 114 111 103 304 115 115 341 294 118 267 379 98 284 314 313 111 32 263 111 108 117 352 289 73 312 299 371 116 114 105 99 116 108 121 32 282 101 97 107 391 341 97 32 346 108 32 102 97 412 329 267 416 99 262 364 102 411 424 359 418 114 99 104 46 34 32 45 32 65 108 259 114 312 69 267 372 101 267

Cnt = 204

StrLen = 291

Compression ratio = 0.7010309278350515

Decompressed output =

"I believe in intuition and inspiration. Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution. It is, strictly speaking, a real factor in scientific research." - Albert Einstein

New dictionary entries:

256 "I

257 I

258 b

259 be

260 el

261 li

262 ie

263 ev

264 ve

265 e

266 i

267 in

268 n

269 in

270 nt

271 tu

272 ui

273 it

274 ti

275 io

276 on

277 n a

278 an

279 nd

280 d

281 ins

282 sp

283 pi

284 ir

285 ra

286 at

287 tio

288 on.

289 .

290 I

291 Im

292 ma

293 ag

294 gi

295 ina

296 ati

297 ion

298 n i

299 is

300 s

301 m

302 mo

303 or

304 re

305 e i

306 im

307 mp

308 po

309 ort

310 ta

311 ant

312 t

313 t

314 th

315 ha

316 an

317 k

318 kn

319 no

320 ow

321 wl

322 le

323 ed

324 dg

325 ge

326 e.

327 . F

328 Fo

329 or

330 kn

331 now

332 wle

333 edg

334 ge

335 is

336 s l

337 lim

338 mi

339 ite

340 ed,

341 ,

342 w

343 wh

344 he

345 er

346 rea

347 as

348 s i

349 ima

350 agi

351 inat

352 tion

353 n e

354 em

355 mb

356 br

357 rac

358 ce

359 es

360 s t

361 the

362 e e

363 en

364 nti

365 ire

366 e w

367 wo

368 orl

369 ld

370 d,

371 , s

372 st

373 tim

374 mu

375 ul

376 la

377 atin

378 ng

379 g

380 p

381 pr

382 ro

383 og

384 gr

385 res

386 ss

387 s,

388 , g

389 giv

390 vi

391 ing

392 g b

393 bi

394 irt

395 th

396 to

397 o

398 e

399 evo

400 ol

401 lu

402 ut

403 tion.

404 . I

405 It

406 t i

407 is,

408 , st

409 tr

410 ri

411 ic

412 ct

413 tl

414 ly

415 y

416 s

417 spe

418 ea

419 ak

420 ki

421 ing,

422 , a

423 a

424 r

425 real

426 l

427 f

428 fa

429 ac

430 cto

431 or i

432 in

433 sc

434 ci

435 ien

436 ntif

437 fi

438 ic

439 re

440 ese

441 ear

442 rc

443 ch

444 h.

445 ."

446 "

447 -

448 -

449 A

450 Al

451 lb

452 ber

453 rt

454 t E

455 Ei

456 ins

457 ste

458 ei

Uncompresed text = "I do not know what I may appear to the world, but to myself I seem to have been only a boy playing on the seashore, and diverting myself in now and then finding a smoother pebble or prettier shell than ordinary, while the great ocean of truth lay all undiscovered before me." - Isaac Newton

\*\*\*\*

Compressed output is 34 73 32 100 111 32 110 111 116 32 107 262 119 32 119 104 97 264 257 109 97 121 32 97 112 112 101 97 114 32 116 260 116 104 101 269 111 114 108 100 44 32 98 117 264 286 32 109 121 115 101 108 102 32 257 305 101 109 285 260 271 118 290 98 101 101 110 32 111 110 108 277 97 297 111 277 112 108 276 105 110 103 323 322 288 290 305 97 115 104 292 101 296 97 110 100 258 105 317 114 116 335 337 303 305 307 32 335 261 111 268 349 351 340 322 102 335 100 357 278 32 115 109 111 263 289 284 281 98 98 108 290 292 32 112 114 101 116 356 101 284 344 306 108 285 271 322 292 373 110 283 121 296 270 105 386 400 290 103 391 272 323 99 282 402 308 116 114 299 104 32 333 327 108 399 117 350 105 115 99 111 354 101 351 319 102 346 302 101 46 34 32 45 309 115 97 97 99 32 78 101 119 286 110

Cnt = 204

StrLen = 291

Compression ratio = 0.7010309278350515

Decompressed output =

"I do not know what I may appear to the world, but to myself I seem to have been only a boy playing on the seashore, and diverting myself in now and then finding a smoother pebble or prettier shell than ordinary, while the great ocean of truth lay all undiscovered before me." - Isaac Newton

New dictionary entries:

256 "I

257 I

258 d

259 do

260 o

261 n

262 no

263 ot

264 t

265 k

266 kn

267 now

268 w

269 w

270 wh

271 ha

272 at

273 t I

274 I m

275 ma

276 ay

277 y

278 a

279 ap

280 pp

281 pe

282 ea

283 ar

284 r

285 t

286 to

287 o t

288 th

289 he

290 e

291 wo

292 or

293 rl

294 ld

295 d,

296 ,

297 b

298 bu

299 ut

300 t t

301 to

302 m

303 my

304 ys

305 se

306 el

307 lf

308 f

309 I

310 I s

311 see

312 em

313 m

314 to

315 o h

316 hav

317 ve

318 e b

319 be

320 ee

321 en

322 n

323 o

324 on

325 nl

326 ly

327 y a

328 a

329 bo

330 oy

331 y p

332 pl

333 la

334 ayi

335 in

336 ng

337 g

338 on

339 n t

340 the

341 e s

342 sea

343 as

344 sh

345 ho

346 ore

347 e,

348 , a

349 an

350 nd

351 d

352 di

353 iv

354 ver

355 rt

356 ti

357 ing

358 g m

359 mys

360 sel

361 lf

362 i

363 in

364 no

365 ow

366 w a

367 and

368 d t

369 then

370 n f

371 fi

372 ind

373 di

374 ing

375 a

376 s

377 sm

378 mo

379 oo

380 oth

381 her

382 r p

383 peb

384 bb

385 bl

386 le

387 e o

388 or

389 p

390 pr

391 re

392 et

393 tt

394 tie

395 er

396 r s

397 she

398 ell

399 l

400 th

401 han

402 n o

403 ord

404 din

405 na

406 ary

407 y,

408 , w

409 whi

410 il

411 le

412 the

413 e g

414 gr

415 rea

416 at

417 oc

418 ce

419 ean

420 n of

421 f t

422 tr

423 ru

424 uth

425 h

426 l

427 lay

428 y al

429 ll

430 l u

431 un

432 ndi

433 is

434 sc

435 co

436 ov

437 vere

438 ed

439 d b

440 bef

441 fo

442 ore

443 me

444 e.

445 ."

446 "

447 -

448 -

449 Is

450 sa

451 aa

452 ac

453 c

454 N

455 Ne

456 ew

457 wt

458 tot