



ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
UIT-HCM

BÁO CÁO VIỆC GIẢI BÀI TOÁN KNAPSACK BẰNG GOOGLE OR TOOLS

Sinh viên thực hiện: Nguyễn Duy Đạt

MSSV: 20520435

Lớp: CS106.M21.KHTN

Giảng viên: Lương Ngọc Hoàng

1. Xử lý dữ liệu

- Link: <https://github.com/likr/kplib>

- Có rất nhiều test trong link trên, việc quét hết các trường hợp khá tốn thời gian nên trong source code ta sẽ dùng random lấy ngẫu nhiên các test từ link.

- Tạo danh sách (list) tên “folder_name” chứa 13 nhóm test cases (00-12) của bài toán:

```
folder_name = ["00Uncorrelated", "01WeaklyCorrelated",  
"02StronglyCorrelated", "03InverseStronglyCorrelated",  
"04AlmostStronglyCorrelated", "05SubsetSum",  
"06UncorrelatedWithSimilarWeights", "07SpannerUncorrelated",  
"08SpannerWeaklyCorrelated", "09SpannerStronglyCorrelated",  
"10MultipleStronglyCorrelated", "11ProfitCeiling", "12Circle"]
```

- Tạo danh sách (list) tên “folder_name_2” chứa các nhóm có kích thước khác nhau 50, 100, 200, ... items từ mỗi folder_name trên :

```
folder_name_2 = ["n00050", "n00100", "n00200", "n00500", "n01000",  
"n02000", "n05000", "n10000"]
```

-Trong mỗi folder_name_2, random 1 số từ 0 -> 1, nếu > 0.5 thì chọn tiếp folder R01000 ngược lại chọn R10000.

-Random 1 số nguyên từ 10->99 trong các folder R01000 và R10000 có dạng s0**.kp (** là số được random ở trên).

2. Lập bảng thống kê

- Chia dữ liệu đầu vào thành 3 nhóm để thống kê:
 - + Dữ liệu nhỏ (small data): kích thước 50, 100, 200, 500 items
 - + Dữ liệu trung bình (medium data): kích thước 1000, 2000 items
 - + Dữ liệu lớn (large data): kích thước 5000, 10000 items
- Bên dưới là thống kê khi thực nghiệm trên laptop cá nhân.

a) Dữ liệu nhỏ

Tên mỗi test case	Giá trị của lời giải	Tổng trọng lượng các items trong lời giải
00Uncorrelated/n00050/R01000/s085.kp	19500	12818
00Uncorrelated/n00100/R01000/s074.kp	42927	25853
00Uncorrelated/n00200/R10000/s023.kp	791932	525855
00Uncorrelated/n00500/R01000/s041.kp	205566	125715
01WeaklyCorrelated/n00050/R10000/s068.kp	143968	130621
01WeaklyCorrelated/n00100/R01000/s072.kp	27537	25037
01WeaklyCorrelated/n00200/R01000/s025.kp	53408	47672
01WeaklyCorrelated/n00500/R10000/s023.kp	1402325	1281298
02StronglyCorrelated/n00050/R01000/s070.kp	16791	13391
02StronglyCorrelated/n00100/R10000/s025.kp	301980	231980
02StronglyCorrelated/n00200/R01000/s023.kp	63421	49421
02StronglyCorrelated/n00500/R01000/s033.kp	158661	123361
03InverseStronglyCorrelated/n00050/R10000/s080.kp	125934	140934
03InverseStronglyCorrelated/n00100/R01000/s068.kp	27696	31096
03InverseStronglyCorrelated/n00200/R01000/s067.kp	53296	59696

03InverseStronglyCorrelated/n00500/R01000/s041.kp	136887	152887
04AlmostStronglyCorrelated/n00050/R10000/s074.kp	166811	132733
04AlmostStronglyCorrelated/n00100/R10000/s072.kp	320330	250193
04AlmostStronglyCorrelated/n00200/R10000/s025.kp	618399	476308
04AlmostStronglyCorrelated/n00500/R01000/s062.kp	156812	121268
05SubsetSum/n00050/R01000/s014.kp	11824	11824
05SubsetSum/n00100/R10000/s076.kp	242434	242434
05SubsetSum/n00200/R01000/s025.kp	47675	47675
05SubsetSum/n00500/R10000/s012.kp	1239316	1239316
06UncorrelatedWithSimilarWeights/n00050/R01000/s064.kp	17718	2401167
06UncorrelatedWithSimilarWeights/n00100/R01000/s076.kp	36530	4902429
06UncorrelatedWithSimilarWeights/n00200/R01000/s062.kp	74081	9904457
06UncorrelatedWithSimilarWeights/n00500/R01000/s038.kp	187781	24712395
07SpannerUncorrelated/n00050/R10000/s037.kp	102300	92565
07SpannerUncorrelated/n00100/R10000/s066.kp	86569	88918
07SpannerUncorrelated/n00200/R01000/s067.kp	39336	32878
07SpannerUncorrelated/n00500/R10000/s055.kp	709426	551848
08SpannerWeaklyCorrelated/n00050/R10000/s081.kp	109210	61640
08SpannerWeaklyCorrelated/n00100/R01000/s016.kp	16984	9898
08SpannerWeaklyCorrelated/n00200/R10000/s068.kp	771730	391590
08SpannerWeaklyCorrelated/n00500/R01000/s096.kp	85102	74282
09SpannerStronglyCorrelated/n00050/R10000/s025.kp	230404	68404
09SpannerStronglyCorrelated/n00100/R10000/s023.kp	491412	224412
09SpannerStronglyCorrelated/n00200/R10000/s093.kp	1026832	358832
09SpannerStronglyCorrelated/n00500/R10000/s022.kp	2633524	693524
10MultipleStronglyCorrelated/n00050/R01000/s033.kp	19608	12108
10MultipleStronglyCorrelated/n00100/R01000/s044.kp	38710	22710
10MultipleStronglyCorrelated/n00200/R01000/s099.kp	81228	50028
10MultipleStronglyCorrelated/n00500/R01000/s090.kp	201190	123090
11ProfitCeiling/n00050/R01000/s056.kp	11199	11205
11ProfitCeiling/n00100/R01000/s099.kp	23367	23377
11ProfitCeiling/n00200/R01000/s060.kp	47118	47140
11ProfitCeiling/n00500/R10000/s090.kp	1229766	1229815
12Circle/n00050/R10000/s020.kp	7777872	116674
12Circle/n00100/R10000/s045.kp	15843263	237661
12Circle/n00200/R10000/s040.kp	31687331	475334
12Circle/n00500/R01000/s021.kp	2624256	124544

b) Dữ liệu trung bình

Tên mỗi test case	Giá trị của lời giải	Tổng trọng lượng các items trong lời giải
00Uncorrelated/n01000/R01000/s031.kp	409620	254916
00Uncorrelated/n02000/R10000/s051.kp	8182011	4960376
01WeaklyCorrelated/n01000/R10000/s099.kp	2674646	2434405
01WeaklyCorrelated/n02000/R10000/s057.kp	5477817	4968368
02StronglyCorrelated/n01000/R10000/s069.kp	3151386	2446386
02StronglyCorrelated/n02000/R10000/s070.kp	6341776	4933776
03InverseStronglyCorrelated/n01000/R10000/s015.kp	2726616	3044616
03InverseStronglyCorrelated/n02000/R10000/s097.kp	5312181	5948181
04AlmostStronglyCorrelated/n01000/R01000/s026.kp	314286	243243
04AlmostStronglyCorrelated/n02000/R01000/s079.kp	632784	491948
05SubsetSum/n01000/R01000/s046.kp	245262	245262
05SubsetSum/n02000/R01000/s099.kp	484289	484289
06UncorrelatedWithSimilarWeights/n01000/R10000/s084.kp	381720	49523476
06UncorrelatedWithSimilarWeights/n02000/R01000/s079.kp	738207	99048931
07SpannerUncorrelated/n01000/R10000/s074.kp	2228464	297157
07SpannerUncorrelated/n02000/R10000/s078.kp	3679280	4176480
08SpannerWeaklyCorrelated/n01000/R10000/s036.kp	5208566	1660778
08SpannerWeaklyCorrelated/n02000/R10000/s085.kp	5097447	2018171
09SpannerStronglyCorrelated/n01000/R10000/s044.kp	4138333	1149333
09SpannerStronglyCorrelated/n02000/R01000/s077.kp	981432	279732
10MultipleStronglyCorrelated/n01000/R10000/s029.kp	3991452	2425452
10MultipleStronglyCorrelated/n02000/R01000/s048.kp	818646	510246
11ProfitCeiling/n01000/R10000/s052.kp	2482929	2483010
11ProfitCeiling/n02000/R01000/s099.kp	484095	484287
12Circle/n01000/R01000/s055.kp	5254949	249393
12Circle/n02000/R01000/s014.kp	10269872	487395

c) Dữ liệu lớn

Tên mỗi test case	Giá trị của lời giải	Tổng trọng lượng các items trong lời giải
00Uncorrelated/n05000/R10000/s056.kp	20336097	12433728
00Uncorrelated/n10000/R01000/s096.kp	4037994	2492322
01WeaklyCorrelated/n05000/R10000/s077.kp	13508865	12197819
01WeaklyCorrelated/n10000/R10000/s079.kp	27374941	24765450
02StronglyCorrelated/n05000/R10000/s092.kp	15769869	12249869
02StronglyCorrelated/n10000/R01000/s045.kp	3186133	2483333
03InverseStronglyCorrelated/n05000/R10000/s015.kp	13388126	14964126
03InverseStronglyCorrelated/n10000/R01000/s080.kp	2658403	2972603
04AlmostStronglyCorrelated/n05000/R01000/s093.kp	1586821	1233099
04AlmostStronglyCorrelated/n10000/R01000/s069.kp	3181579	2478223
05SubsetSum/n05000/R10000/s084.kp	12532088	12532088
05SubsetSum/n10000/R01000/s013.kp	2473178	2473178
06UncorrelatedWithSimilarWeights/n05000/R01000/s060.kp	1854891	247625357
06UncorrelatedWithSimilarWeights/n10000/R10000/s068.kp	3764471	495242987
07SpannerUncorrelated/n05000/R10000/s024.kp	12848751	7192039
07SpannerUncorrelated/n10000/R10000/s060.kp	12419433	13267593
08SpannerWeaklyCorrelated/n05000/R01000/s049.kp	826800	294150
08SpannerWeaklyCorrelated/n10000/R01000/s018.kp	1387603	1032132
09SpannerStronglyCorrelated/n05000/R01000/s098.kp	2451344	727044
09SpannerStronglyCorrelated/n10000/R01000/s020.kp	4996194	1946494
10MultipleStronglyCorrelated/n05000/R10000/s017.kp	20137196	12422196
10MultipleStronglyCorrelated/n10000/R10000/s089.kp	40278006	24729006
11ProfitCeiling/n05000/R10000/s068.kp	12485235	12485707
11ProfitCeiling/n10000/R10000/s048.kp	25063698	25064633
12Circle/n05000/R01000/s031.kp	26392812	1252569
12Circle/n10000/R01000/s052.kp	51942031	2465103

3. Nhận xét

- Với dữ liệu nhỏ (small data) và dữ liệu trung bình (medium data), Google OR Tools có thể đưa ra được lời giải tối ưu. Bên cạnh đó vẫn có 1 số trường hợp ngoại lệ bị tràn RAM.

- Với dữ liệu lớn (large data), Google OR Tools trong các trường hợp bị tràn RAM mặc dù vẫn đưa ra được kết quả như bảng thống kê trên.

- Trong 13 nhóm test cases: (quan sát trong folder output)

- + Nhóm khó giải là 12Circle, 11ProfitCeiling, 09MultipleStronglyCorrelated, 08SpannerWeaklyCorrelated có nhiều test item có thời gian thực thi > 200 giây.

- + Nhóm dễ giải là 00Uncorelated và 01WeaklyCorrelated khi các test item có thời gian thực thi < 1 giây.