

# Apprentissage et résultats

Clément Legrand

July 6, 2018

# Description

## Base de départ

Les solutions données par CW.

- Tirage au sort de  $N$  triplets  $(\lambda, \mu, \nu)$ ;
- Calcul des solutions pour tout triplet  $(\lambda, \mu, \nu)$ .

## Base d'apprentissage

On peut ne garder qu'une partie de la base générée pour apprendre

- On garde  $x\%$  des meilleures solutions (quantité privilégiée,  $\text{Quan}_x$ );
- On garde les solutions qui ont un coût inférieur à  $c_{min} + (c_{max} - c_{min}) \frac{x}{100}$  (qualité privilégiée,  $\text{Qual}_x$ ).
- On choisit d'utiliser toute la base générée pour apprendre (Tout)

# Protocole

## Protocole

- Génération d'un échantillon de taille  $N_{ech}$
- Calcul de la base d'apprentissage
- On initialise une matrice MAT de taille  $n^2$
- Pour chaque arête (a,b) on incrémente la valeur MAT[a][b] (si  $a > b$ , on commence par échanger a et b)
- Comparaison arêtes obtenues et optimales.

## Choix des arêtes

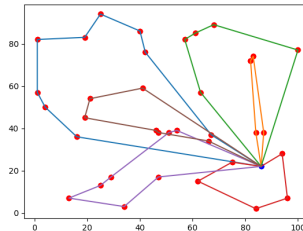
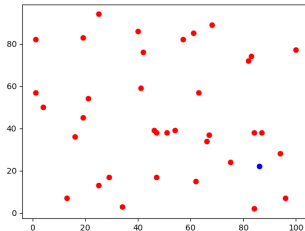
- On conserve (a,b) si MAT[a][b] dépasse une certaine valeur (Seuil);
- On conserve les k premières arêtes en triant selon les valeurs contenues dans MAT (Rang).

# Instance test

3 instances ont été choisies pour réaliser ces tests: A-n37-k06, A-n65-k09 et P-n101-k04.

La solution employée pour comparer les résultats est celle de la littérature.

La meilleure solution comporte 42 arêtes.



Pour chaque test on effectue 5 itérations.

Temps de calcul: 2 s (50), 4 s (100), 20 s (500), 44 s (1000).

## Résultats A-n37-k06, critère Seuil

	Quan <sub>10</sub>				Qual <sub>10</sub>				Tout			
	Seuil	Arêtes	Corr	Prop	Seuil	Arêtes	Corr	Prop	Seuil	Arêtes	Corr	Prop
50	3	34	21	0.5	11	33	21	0.50	25	23	15	0.35
	4	23	14	0.33	17	17	12	0.28	38	10	7	0.16
100	5	30	21	0.5	15	31	23	0.55	50	24	17	0.40
	8	16	15	0.36	23	17	14	0.33	75	6	6	0.14
500	25	32	24	0.57	58	31	22	0.52	250	22	15	0.36
	38	15	14	0.33	88	20	16	0.38	375	7	7	0.18
Comple	400	33	24	0.57	732	30	23	0.55	4000	25	16	0.38
	600	15	14	0.33	1097	18	16	0.38	6000	9	6	0.14

- Taille de l'échantillon ne semble pas avoir d'influence sur les résultats (*prop* reste semblable quel que soit la taille de l'échantillon).
- Avec base *Tout* : valeurs de *prop* plus basses → pas la peine d'utiliser tout l'échantillon.

Remarque : Base Quan<sub>10</sub> trop petite avec échantillon 50 ou 100.

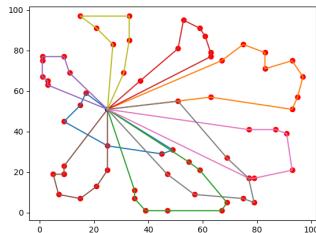
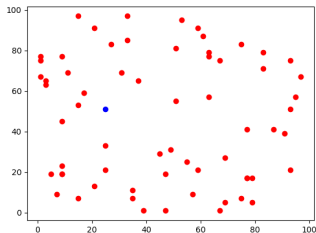
## Résultats A-n37-k06, critère Rang

	Quan <sub>10</sub>			Qual <sub>10</sub>			Tout		
	Rang	Corr	Prop	Rang	Corr	Prop	Rang	Corr	Prop
50	10	6	0.14	10	6	0.14	10	7	0.16
	20	13	0.31	20	13	0.32	20	13	0.31
	18	12	0.28	18	13	0.3	18	12	0.28
100	10	9	0.21	10	9	0.21	10	10	0.24
	20	16	0.38	20	16	0.38	20	15	0.36
	18	13	0.3	18	13	0.3	18	12	0.29
500	10	9	0.21	10	10	0.24	10	9	0.21
	20	16	0.38	20	16	0.38	20	15	0.36
	18	13	0.3	18	13	0.3	18	12	0.28
Complet	10	8	0.19	10	9	0.21	10	7	0.17
	20	14	0.33	20	14	0.33	20	14	0.33
	18	12	0.29	18	12	0.29	18	12	0.29

Les 3 bases fournissent des valeurs *prop* similaires.

# Instance test

La solution employée pour comparer les résultats est celle de la littérature.  
La meilleure solution comporte 73 arêtes.



Pour chaque test on effectue 5 itérations.

Temps de calcul: 4 s (50), 8 s (100), 42 s (500), 84 s (1000).

## Résultats A-n65-k09, critère Seuil

	Quan <sub>10</sub>				Qual <sub>10</sub>				Tout			
	Seuil	Arêtes	Corr	Prop	Seuil	Arêtes	Corr	Prop	Seuil	Arêtes	Corr	Prop
50	3	73	43	0.59	10	64	44	0.60	25	40	31	0.43
	4	61	40	0.55	15	39	29	0.40	38	14	9	0.13
100	5	70	44	0.6	22	58	42	0.58	50	43	33	0.45
	8	63	41	0.56	33	36	28	0.39	75	15	10	0.14
500	25	71	43	0.59	111	56	41	0.56	250	45	35	0.48
	38	60	40	0.55	167	35	28	0.39	375	14	9	0.13
Complet	400	62	41	0.56	1005	56	40	0.55	4000	45	35	0.48
	600	15	14	0.33	1508	35	28	0.39	6000	13	9	0.12

Si trop d'arêtes renvoyées → Solutions infaisables ? (futurs tests)



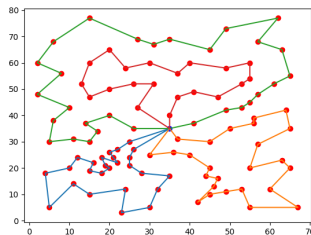
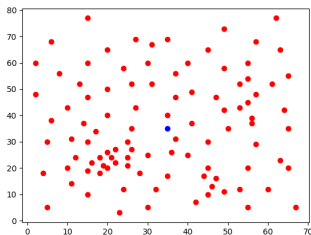
## Résultats A-n65-k09, critère Rang

	Quan <sub>10</sub>			Qual <sub>10</sub>			Tout		
	Rang	Corr	Prop	Rang	Corr	Prop	Rang	Corr	Prop
50	10	6	0.08	10	7	0.1	10	7	0.1
	20	14	0.2	20	15	0.21	20	14	0.19
	33	23	0.32	33	26	0.36	33	24	0.33
100	10	6	0.08	10	7	0.1	10	7	0.1
	20	16	0.22	20	16	0.22	20	14	0.19
	33	26	0.36	33	26	0.36	33	25	0.34
500	10	7	0.1	10	7	0.1	10	6	0.08
	20	17	0.23	20	15	0.21	20	13	0.18
	33	27	0.37	33	26	0.36	33	25	0.34
Complet	10	7	0.1	10	7	0.1	10	6	0.08
	20	17	0.23	20	17	0.23	20	13	0.18
	33	27	0.37	33	27	0.37	33	25	0.34

De nouveau les 3 bases renvoient des résultats similaires.

## Instance test

La solution employée pour comparer les résultats est celle de la littérature.  
La meilleure solution comporte 104 arêtes.



Pour chaque test on effectue 5 itérations.

Temps de calcul: 38 s (50), 75 s (100), 375 s (500), 1060 s (8000).

## Résultats P-n101-k04, critère Seuil

	Quan <sub>10</sub>				Qual <sub>10</sub>				Tout			
	Seuil	Arêtes	Corr	Prop	Seuil	Arêtes	Corr	Prop	Seuil	Arêtes	Corr	Prop
50	3	93	65	0.62	5	83	66	0.64	25	71	61	0.59
	4	54	44	0.42	8	42	37	0.36	38	24	21	0.20
100	5	80	66	0.64	9	79	66	0.63	50	72	62	0.60
	8	45	41	0.40	14	42	39	0.38	75	24	22	0.21
500	25	83	69	0.67	44	81	68	0.66	250	72	63	0.60
	38	43	39	0.38	67	39	36	0.35	375	22	20	0.19
Complet	400	87	73	0.7	411	85	71	0.68	4000	70	60	0.58
	600	42	39	0.38	616	41	38	0.37	6000	23	21	0.2

Plus la taille de l'instance augmente, et plus la proportion d'arêtes optimales renvoyées présentes dans la solution optimale est grande.

## Résultats P-n101-k04, critère Rang

	Quan <sub>10</sub>			Qual <sub>10</sub>			Tout		
	Rang	Corr	Prop	Rang	Corr	Prop	Rang	Corr	Prop
50	10	8	0.08	10	8	0.08	10	8	0.08
	20	18	0.17	20	17	0.16	20	18	0.17
	50	43	0.41	50	44	0.43	50	44	0.43
100	10	8	0.08	10	8	0.08	10	8	0.08
	20	18	0.17	20	18	0.17	20	18	0.17
	50	46	0.44	50	46	0.44	50	46	0.44
500	10	8	0.08	10	8	0.08	10	8	0.08
	20	18	0.17	20	18	0.17	20	18	0.17
	50	46	0.44	50	46	0.44	50	46	0.44
Complet	10	8	0.08	10	8	0.08	10	8	0.08
	20	18	0.17	20	18	0.17	20	18	0.17
	50	46	0.44	50	46	0.44	50	46	0.44

Il faut choisir un rang dépendant de la taille de l'instance (rangs fixés à 10 ou 20 ne revoient plus de bons résultats).

# Choix de l'échantillon et de la base

D'après les résultats précédents :

- Taille échantillon : 50 (aussi efficace que les tailles plus grandes, et plus rapide)
- Base d'apprentissage : Qual (Quasi trop petite pour 50, Tout pas intéressante)

Que choisir comme critère pour extraire les arêtes ?

## Nouveaux résultats pour les critères

4 critères : Rang =  $n/2$  ou Rang =  $n$  et Seuil =  $S_{lb}/2$  ou  $S_{lb}/3$

A-n37-k06				A-n65-k09				P-n101-k04			
Critère	Arêtes	Corr.	Infais.	Critère	Arêtes	Corr.	Infais.	Critère	Arêtes	Corr.	Infais.
Rg = 18	18	12	0 (0)	Rg = 32	32	25	0 (0)	Rg = 50	50	45	0 (0)
Rg = 36	36	22	1 (0)	Rg = 64	64	44	2 (1)	Rg = 100	100	67	11 (5)
Se = 8	45	22	5 (4)	Se = 9	81	50	17 (11)	Se = 5	121	56	30 (24)
Se = 12	31	20	0 (0)	Se = 13	59	42	0 (0)	Se = 8	80	62	2 (1)

- Critère Rg =  $n/2$  plus précis mais moins d'arêtes
- Critère Se =  $S_{lb}/2$  moins mais plus d'arêtes
- Les autres critères demandent d'éliminer trop d'arêtes

# Algorithme d'optimisation ( $H_c$ )

---

---

```
1  $Sol \leftarrow CW(\lambda, \mu, \nu)$ 
2  $NewSol \leftarrow Sol$ 
3 while La dernière amélioration date de moins de 10 sec do
4   Calcul de la pire arête
5    $NewSol \leftarrow EjectionChain_{FI-RD}$ 
6    $NewSol \leftarrow LinKernighan_{BI-O}$ 
7    $NewSol \leftarrow CrossExchange_{FI-RD}$ 
8    $NewSol \leftarrow LinKernighan_{BI-O}$ 
9   if  $cost(NewSol) < cost(Sol)$  then
10     $Sol \leftarrow NewSol$ 
11   if Pas d'amélioration depuis  $n/2$  itérations then
12     $NewSol \leftarrow Sol$ 
13 return  $Sol$ 
```

---

# Learning Heuristic

---

```
1  $(\lambda^*, \mu^*, \nu^*), Init \leftarrow Apprentissage()$ 
2  $newBase \leftarrow []$ 
3 for  $i \leftarrow 1$  to 10 do
4     if  $i = 1$  then
5         for  $j \leftarrow 1$  to 10 do
6              $Sol \leftarrow H_c(Init, I, D, \lambda^*, \mu^*, \nu^*)$ 
7              $newBase \leftarrow newBase \cup Sol$ 
8     else
9         Déterminer  $Init$  avec les connaissances de  $newBase$ 
10         $(\lambda^*, \mu^*, \nu^*), Init \leftarrow Apprentissage(Init)$ 
11        for  $j \leftarrow 1$  to 10 do
12             $Sol \leftarrow H_c(Init, I, D, \lambda^*, \mu^*, \nu^*)$ 
13             $newBase \leftarrow newBase \cup Sol$ 
14 return La meilleure solution
```

---



# Premiers résultats

Résultats pour les coûts obtenus

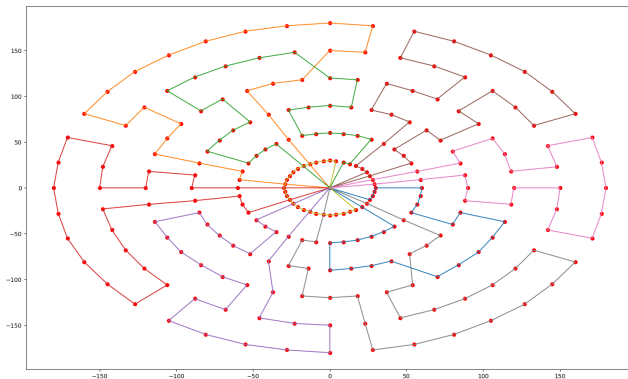
	A3706	A6509	P10104
Best	952	1182	692
$n/2$	950 - 954	1187 - 1202	697 - 714
$S_{lb}/2$	957 - 968	1205 - 1232	697 - 718

Résultats pour le temps d'exécution (en sec)

Instance	A-n37-k06	A-n65-k09	P-n101-k04
Temps (app - total)	8 - 707	25 - 814	98 - 1110

# Nouveau meilleur résultat

Pour l'instance Golden-01, nouvelle solution trouvée:



Coût de 5563 au lieu de 5623