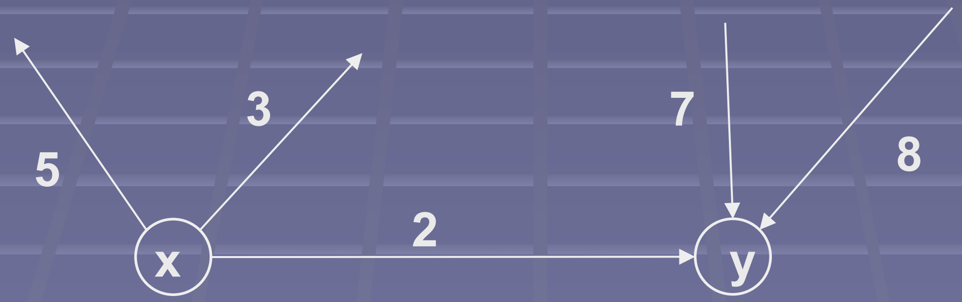
Heuristiques gloutonnes pour le PVC

- Nous avons déja vu l'heuristique du **plus proche voisin** (**PPV**), la plus naïve. Elle s'implémente en $O(N^2)$.
- Pour le Δ-PVC sa performance relative au pire est $\log_2 N$. C'est franchement mauvais : pour 1024 sommets elle peut trouver un coût 10 fois supérieur au coût optimal...
- Sa performance moyenne est meilleure.

Heuristiques gloutonnes pour le PVC

- Une approche moins naïve consiste à ajouter à chaque étape non pas l'arête de coût minimal au départ de x, mais celle de **regret maximal**. C'est une **méthode à pénalité**.



- [x,y] a pour coût 2 et pour regret (ou pénalité) 3+7=10. $R_{[x,y]} = \min_{s \neq y}(C_{xs}) + \min_{s \neq x}(C_{sy})$
- Les méthodes par pénalité sont meilleures en moyenne car moins locales, mais en $O(n^3)$...

Méthode à pénalité pour le PVC

- Reprenons l'exemple précédent :

	A	B	C	D	E
A	-	3	4	5	4
B	3	-	2	2	1
C	4	2	-	1	2
D	5	2	1	-	3
E	4	1	2	3	-

Méthode à pénalité pour le PVC

- Reprenons l'exemple précédent :

	A	B	C	D	E
A	-	3(5)	4(4)	5(4)	4(4)
B	3(5)	-	2(2)	2(2)	1(4)
C	4(4)	2(2)	-	1(4)	2(2)
D	5(3)	2(2)	1(4)	-	3(2)
E	4(4)	1(4)	2(3)	3(2)	-

Méthode à pénalité pour le PVC

- Reprenons l'exemple précédent :

	A	B	C	D	E
A	-	3(5)	4(4)	5(4)	4(4)
B	3(5)	-	2(2)	2(2)	1(4)
C	4(4)	2(2)	-	1(4)	2(2)
D	5(3)	2(2)	1(4)	-	3(2)
E	4(4)	1(4)	2(3)	3(2)	-

- A-B
- coût=3

Méthode à pénalité pour le PVC

- Reprenons l'exemple précédent :

	A	B	C	D	E
A	-	-	-	-	-
B	-	-	2(2)	2(2)	**1(4)**
C	4	-	-	1	2
D	5	-	1	-	3
E	4	-	2	3	-

- A-B-E
- coût=3+1

Méthode à pénalité pour le PVC

- Reprenons l'exemple précédent :

	A	B	C	D	E
A	-	-	-	-	-
B	-	-	-	-	-
C	4	-	-	1	-
D	5	-	1	-	-
E	-	-	**2(4)**	3(3)	-

- A-B-E-C
- coût=3+1+2

Méthode à pénalité pour le PVC

- Reprenons l’exemple précédent :

	A	B	C	D	E
A	-	-	-	-	-
B	-	-	-	-	-
C	-	-	-	1	-
D	5	-	-	-	-
E	-	-	-	-	-

- A-B-E-C-D-A : cycle hamiltonien
- coût=3+1+2+1+5=12 (optimal ici...)

Méthodes par insertion pour le PVC

- On part d'un cycle U trivial sur un sommet arbitraire.
- A chaque itération on choisit un sommet libre k, et on cherche la position d'insertion entre deux sommets consécutifs i,j de U, qui minimise l'augmentation de coût :

$$\Delta C = C_{ik} + C_{kj} - C_{ij}$$

- On change U en un cycle ayant un sommet supplémentaire en supprimant l'arête [i,j] et en ajoutant [i,k] et [k,j].
- On poursuit tant qu'il reste un sommet libre.

Méthodes par insertion pour le PVC

Selon le choix du sommet libre on a différentes heuristiques :

- Plus proche insertion (PPI) :

  Le sommet choisi est le plus proche de U

  (où dist(x,U)=min $\{C_{xy}$: y sommet de U$\}$.)

- Plus lointaine insertion (PLI) :

  Le sommet choisi est le plus éloigné de U

- Meilleure insertion (MI) :

  Le sommet choisi est celui qui donne la plus faible augmentation de coût.

PPI

	A	B	C	D	E
A	-	3	4	5	4
B	3	-	2	2	1
C	4	2	-	1	2
D	5	2	1	-	3
E	4	1	2	3	-

A

PLI

	A	B	C	D	E
A	-	3	4	5	4
B	3	-	2	2	1
C	4	2	-	1	2
D	5	2	1	-	3
E	4	1	2	3	-

A

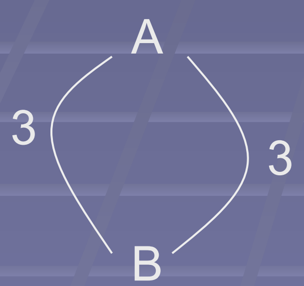
MI

	A	B	C	D	E
A	-	3	4	5	4
B	3	-	2	2	1
C	4	2	-	1	2
D	5	2	1	-	3
E	4	1	2	3	-

A

On choisit (arbitrairement ici...) A comme sommet initial.

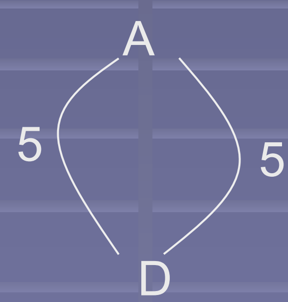
PPI	*A*	**B**	C	D	E
A	-	**3**	4	5	4
B	**3**	-	2	2	1
C	4	2	-	1	2
D	5	2	1	-	3
E	4	1	2	3	-



Le plus proche est B

(distance=3) Δ=6

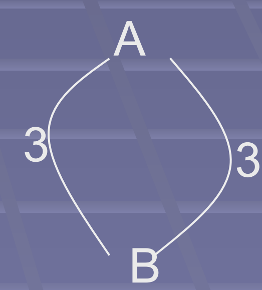
PLI	*A*	B	C	**D**	E
A	-	3	4	**5**	4
B	3	-	2	2	1
C	4	2	-	1	2
D	**5**	2	1	-	3
E	4	1	2	3	-



Le plus éloigné est D

(distance=5) Δ=10

MI	*A*	**B**	C	D	E
A	-	**3**	4	5	4
B	**3**	-	2	2	1
C	4	2	-	1	2
D	5	2	1	-	3
E	4	1	2	3	-

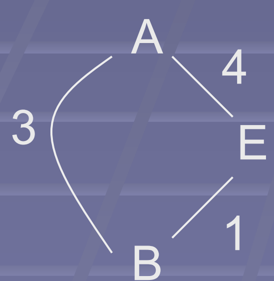


Le meilleur est B

Δ=6

PPI

	A	B	C	D	**E**
A	-	-	4	5	**4**
B	-	-	2	2	**1**
C	4	2	-	1	2
D	5	2	1	-	3
E	**4**	**1**	2	3	-

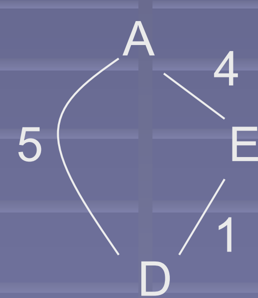


Le plus proche est E

(distance 1) Δ=2

PLI

	A	B	C	D	**E**
A	-	3	4	-	**4**
B	3	-	2	2	1
C	4	2	-	1	2
D	-	2	1	-	**3**
E	**4**	1	2	**3**	-

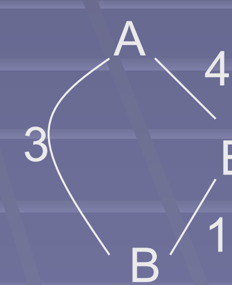


Le plus éloigné est D

(distance 3) Δ=0

MI

	A	B	C	D	**E**
A	-	-	4	5	**4**
B	-	-	2	2	**1**
C	4	2	-	1	2
D	5	2	1	-	3
E	**4**	**1**	2	3	-

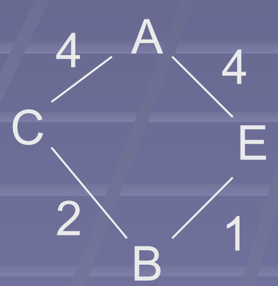


Le meilleur est E

(Δ=2)

PPI

	A	B	**C**	D	E
A	-	-	**4**	5	-
B	-	-	**2**	2	-
C	**4**	**2**	-	1	**2**
D	5	2	1	-	3
E	-	-	**2**	3	-

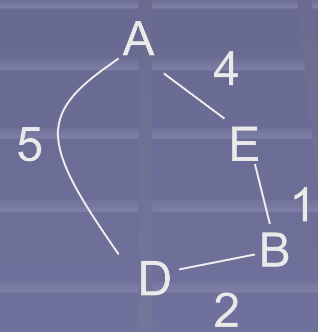


Le plus proche est C

(et D, distance 2) Δ=3

PLI

	A	**B**	C	D	E
A	-	**3**	4	-	-
B	**3**	-	2	**2**	**1**
C	4	2	-	1	2
D	-	**2**	1	-	-
E	-	**1**	2	-	-

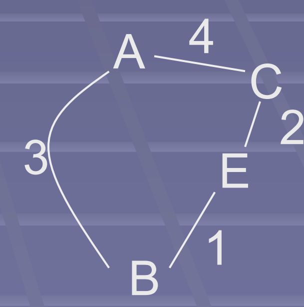


Le plus éloigné est B

(et C, distance 1) Δ=2

MI

	A	B	**C**	D	E
A	-	-	**4**	5	-
B	-	-	**2**	2	-
C	**4**	**2**	-	1	**2**
D	5	2	1	-	3
E	-	-	**2**	3	-

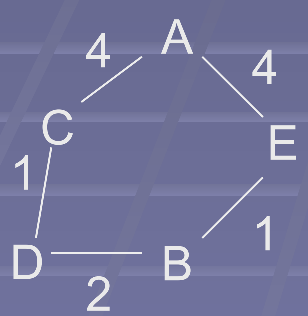


Le meilleur est C

(Δ=2)

PPI

	A	B	C	D	E
A	-	-	-	5	-
B	-	-	-	2	-
C	-	-	-	1	-
D	5	2	1	-	3
E	-	-	-	3	-

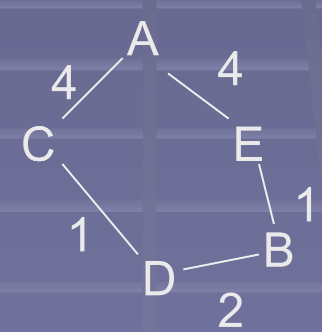


Il ne reste que D

Δ =1

PLI

	A	B	C	D	E
A	-	-	4	-	-
B	-	-	2	-	-
C	4	2	-	1	2
D	-	-	1	-	-
E	-	-	2	-	-

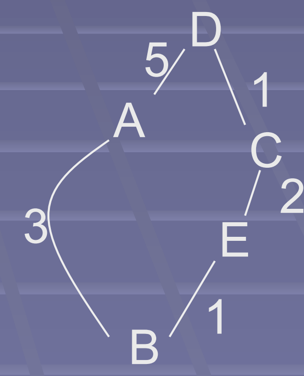


Il ne reste que C

Δ=0

MI

	A	B	C	D	E
A	-	-	-	5	-
B	-	-	-	2	-
C	-	-	-	1	-
D	5	2	1	-	3
E	-	-	-	3	-

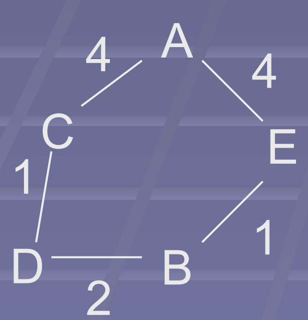


Il ne reste que D

Δ=2

PPI

	A	B	C	D	E
A	-	-	-	5	-
B	-	-	-	2	-
C	-	-	-	1	-
D	5	2	1	-	3
E	-	-	-	3	-

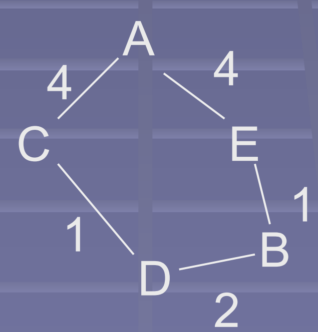


A-E-B-D-C-A, coût 12

cycle hamiltonien

PLI

	A	B	C	D	E
A	-	-	4	-	-
B	-	-	2	-	-
C	4	2	-	1	2
D	-	-	1	-	-
E	-	-	2	-	-

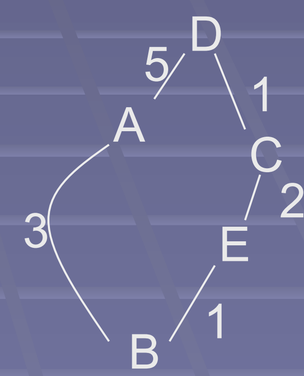


A-E-B-D-C-A, coût 12

cycle hamiltonien

MI

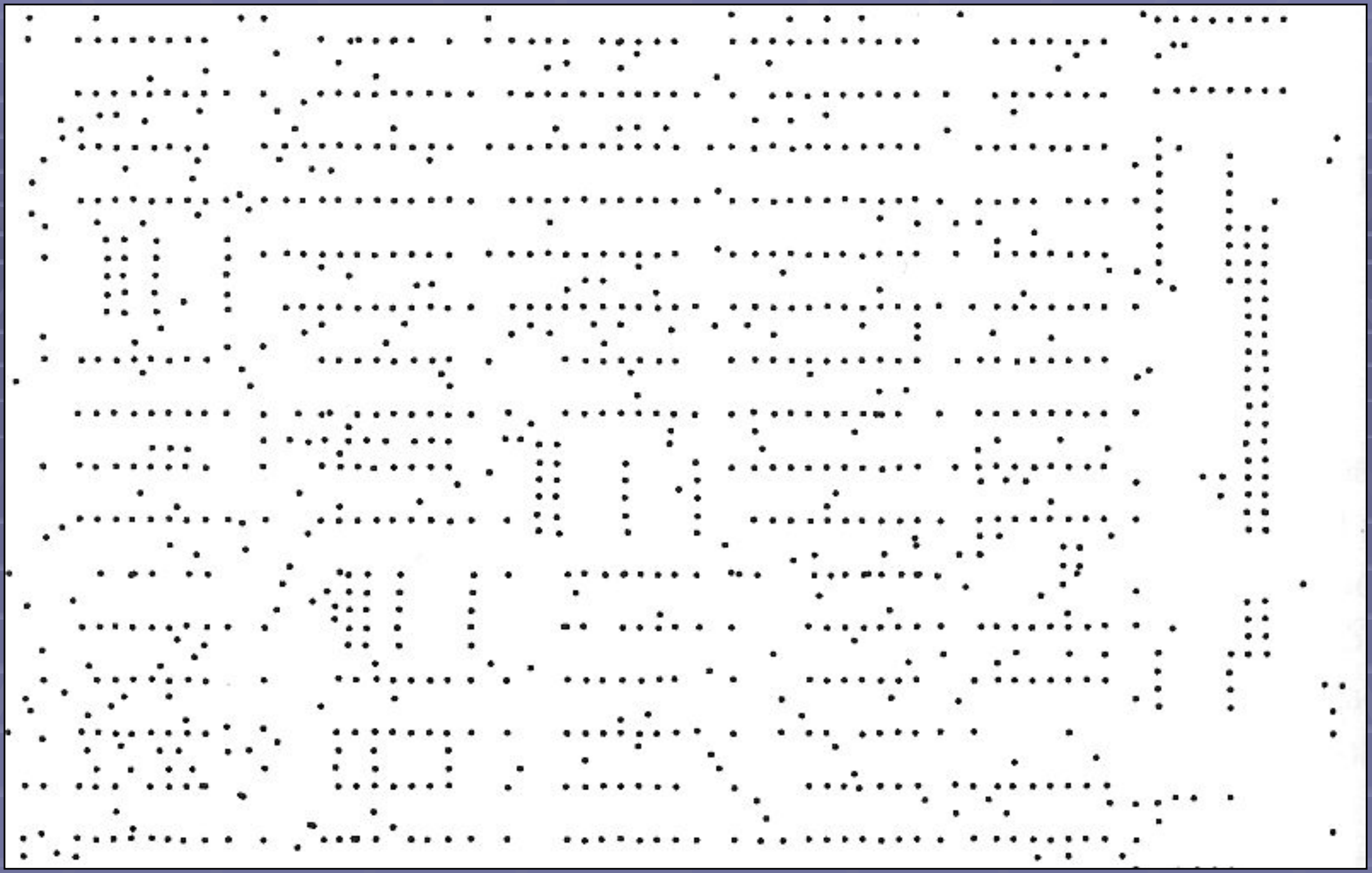
	A	B	C	D	E
A	-	-	-	5	-
B	-	-	-	2	-
C	-	-	-	1	-
D	5	2	1	-	3
E	-	-	-	3	-



A-D-C-E-B-A, coût 12

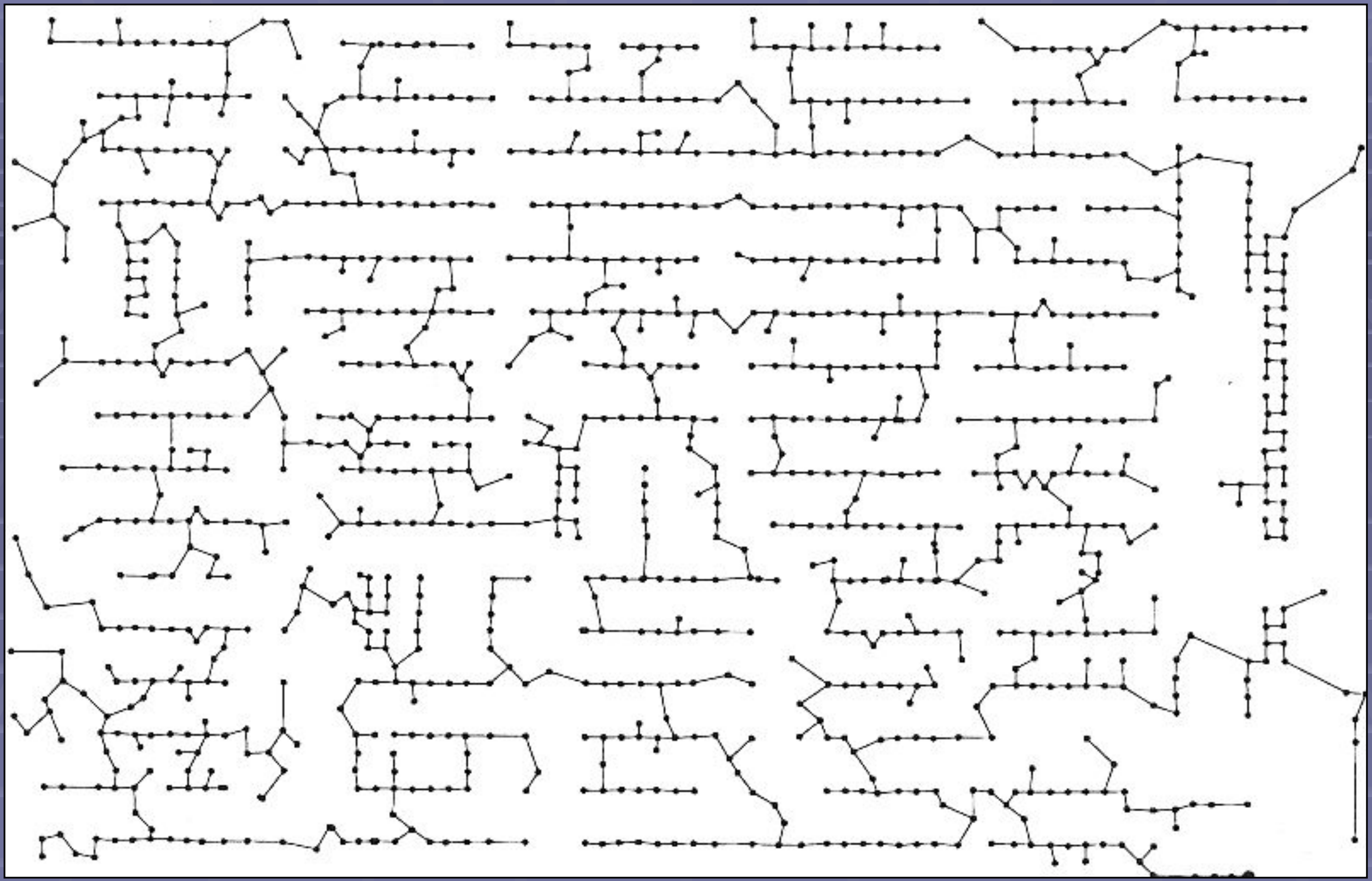
cycle hamiltonien

- PLI et PPI s'implémentent en $O(n^2)$; MI en $O(n^3)$. On prouve (à priori) que leur PRP pour le Δ-PVC est au pire 2.
- Leur performance est cependant meilleure en moyenne. En moyenne MI est meilleure que PPI qui est meilleure que PLI, dont la performance moyenne (statistique) est 1.16 (testée sur TSPLIB).
- Pour des PVC euclidiens, étonnement (?..) la meilleure est PLI. Ce qui montre la nécessité d'évaluer les heuristiques...



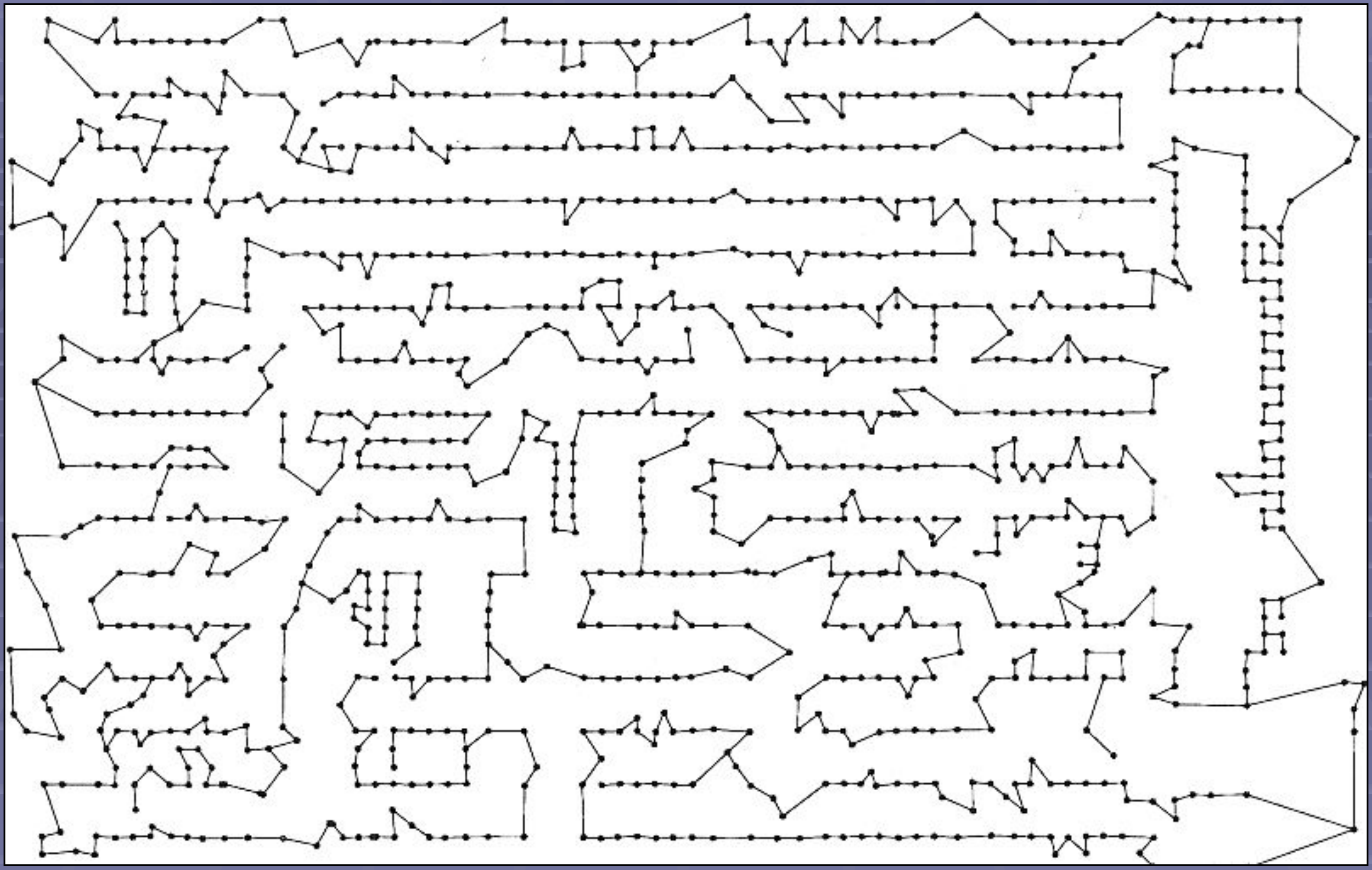
Un PVC euclidien de 1173 sommets.

(Le record (août 1994) est de 7397 sommets.)



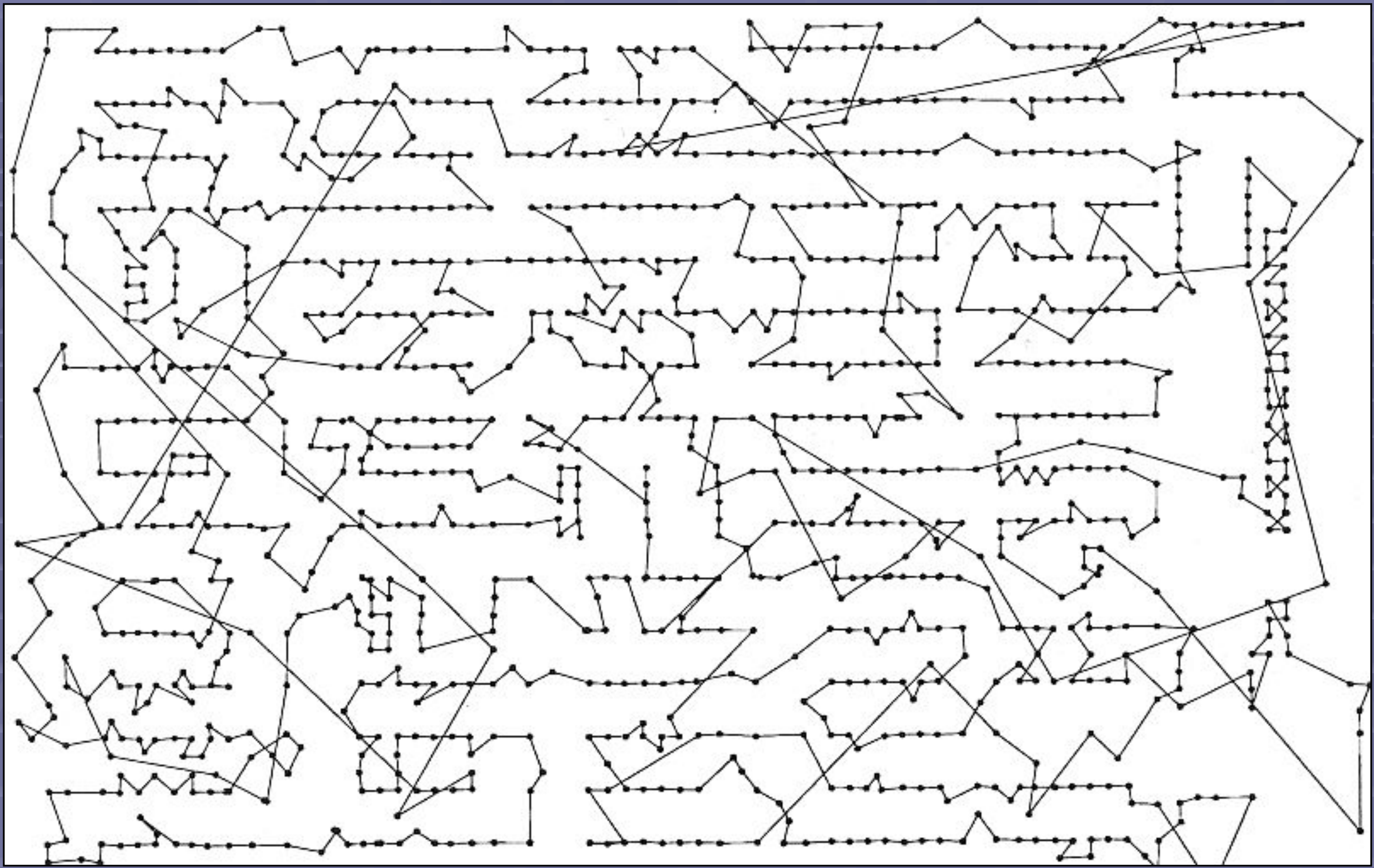
Borne minimale : arbre recouvrant de coût minimal

coût = 51488



Borne minimale plus fine : borne de Held-Karp

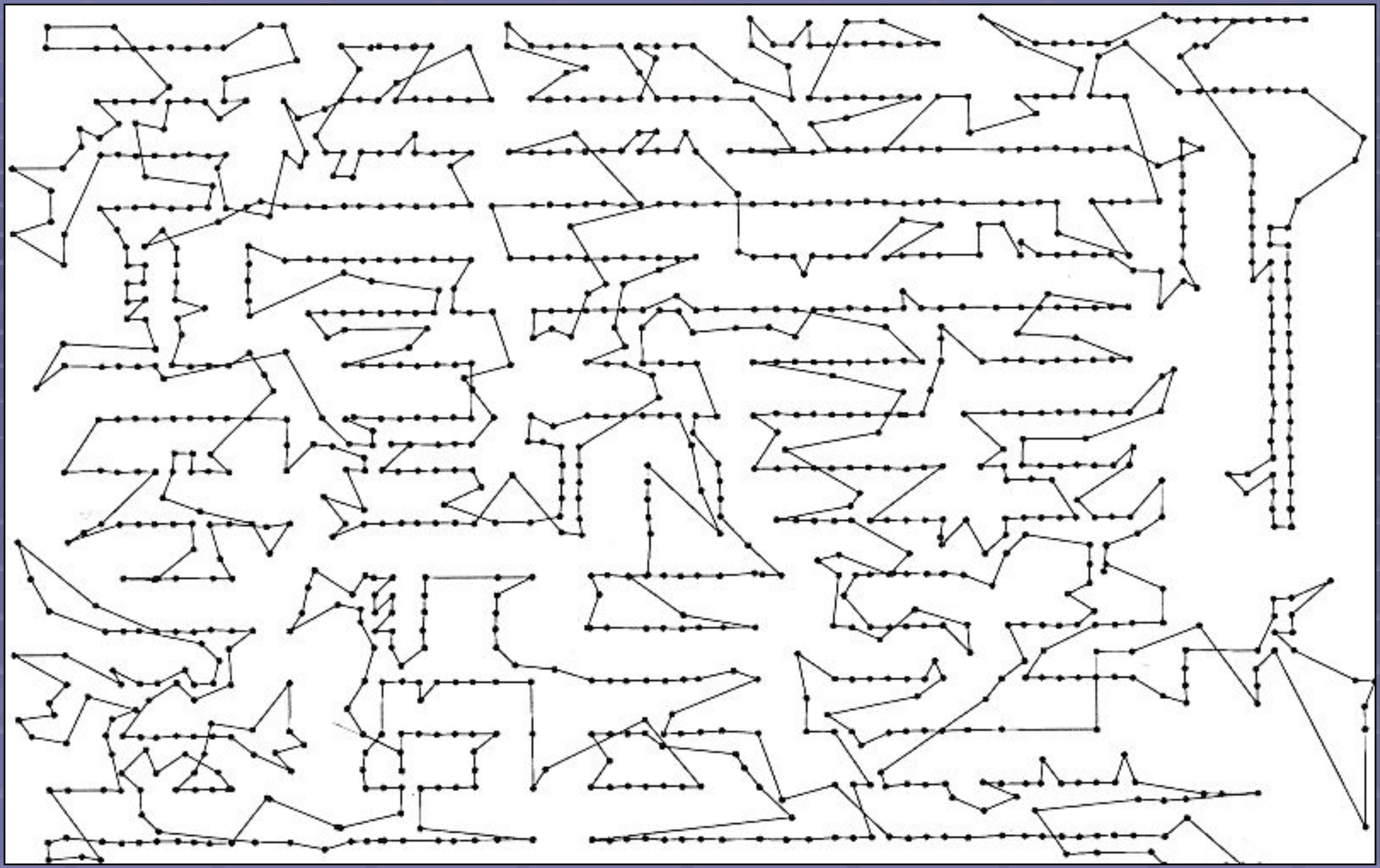
coût = 56349



Heuristique : Plus Proche Voisin

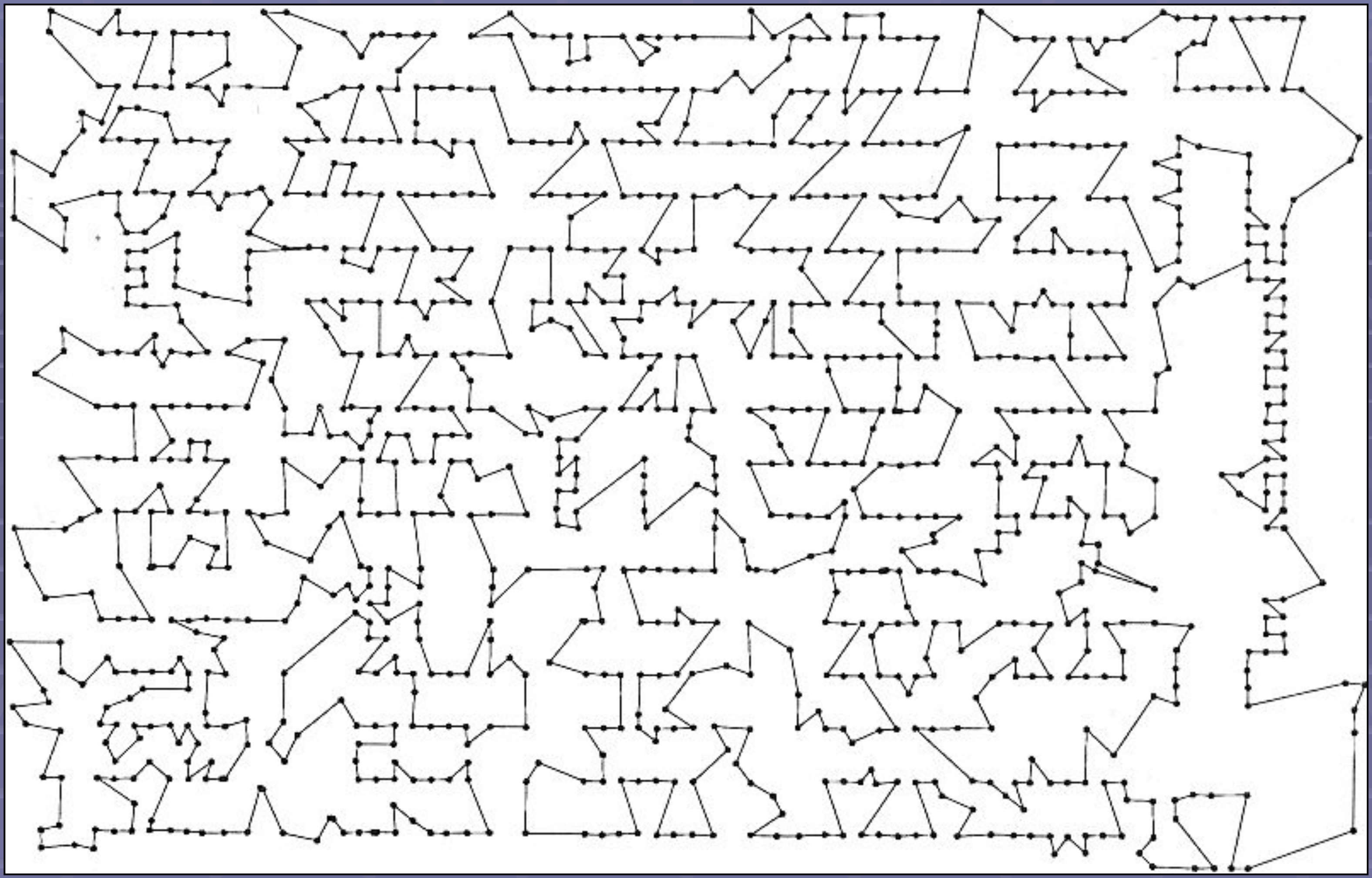
coût = 67822 ; erreur au plus 20%

Remarquer les nombreux « longs trajets »



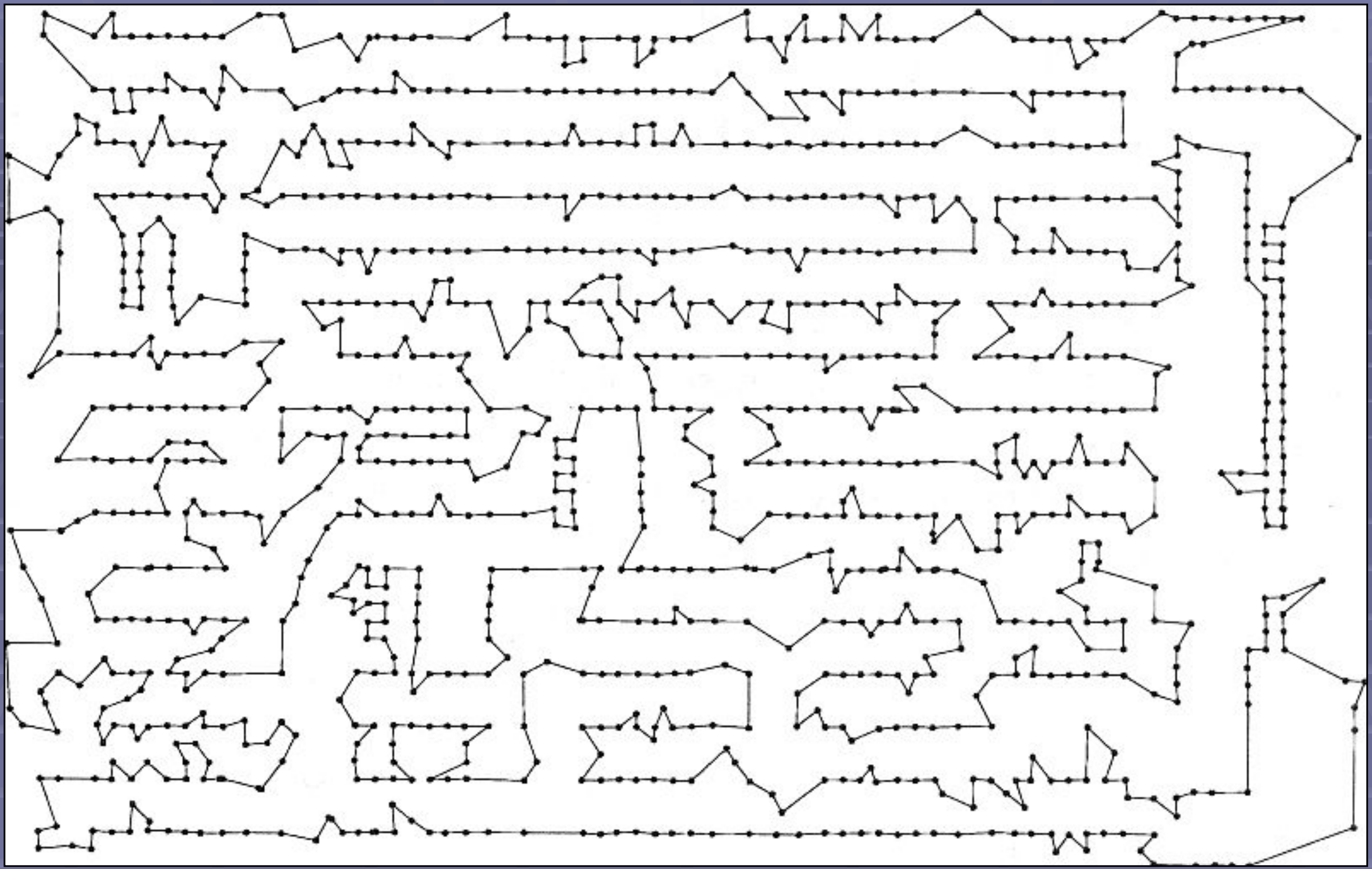
Heuristique : Plus Proche Insertion

coût = 72337 ; erreur au plus 28%



Heuristique : Plus Lointaine Insertion

coût = 65980 ; erreur au plus 17%



Meilleure heuristique connue : Lin-Kernighan
(Recherche locale, d'erreur 1% sur TSPLIB)

coût = 56892 ; erreur au plus 1%

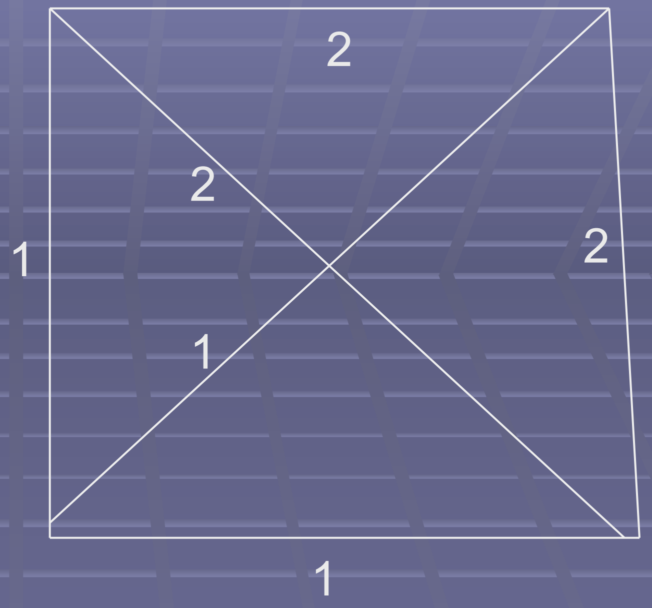
Heuristique de Christofide

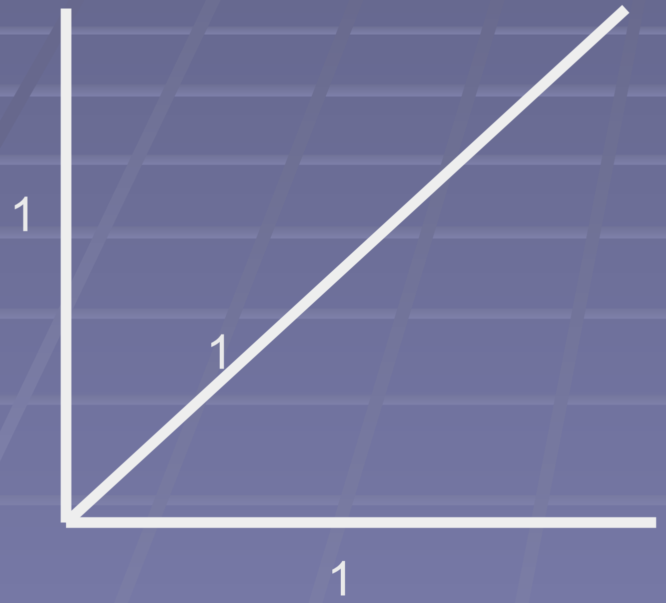
C'est l'heuristique de meilleure **PRP=1,5** connue pour le Δ - PVC !

Heuristique de Christofide

- Soit G graphe complet valué pour le Δ-PVC.
- Déterminer un arbre recouvrant T de coût minimal.
- Soit I les sommets de T de degré impair. Déterminer dans G un couplage C parfait de I de coût minimal.
- Construire un graphe $T_1 \supset T$, avec $T_1 = T \sqcup C$ (ajouter les arêtes de C qui ne sont pas dans T, et dupliquer celles qui sont déjà dans T). Tous les sommets de T_1 sont de degré pair.
- Pour chaque sommet u de degré ≥ 4 dans T_1 supprimer des arêtes [v,u] et [u,w] de T_1 et ajouter [v,w] à T_1 en gardant le graphe connexe, (et tant qu'à faire, de surcoût minimal). (« shortcut »)
- Dès qu'il n'y a plus de sommet de degré $\geq$ 4, on a construit un cycle hamiltonien.

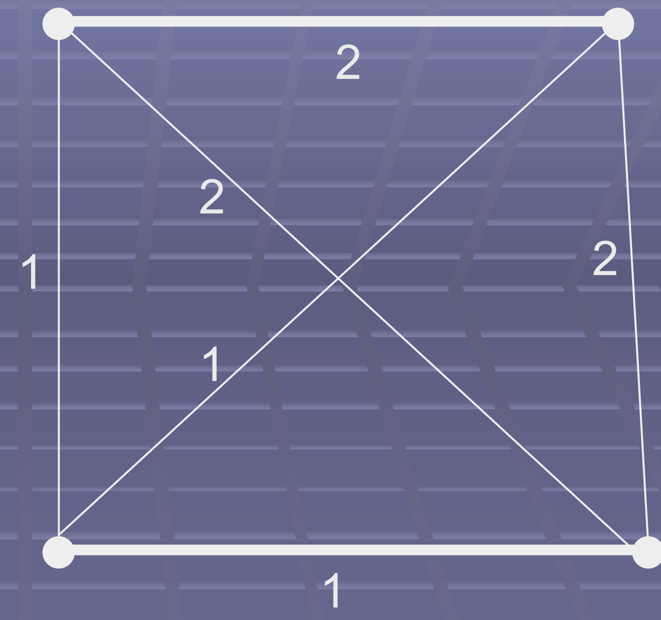
- **Exemple** ; graphe G :

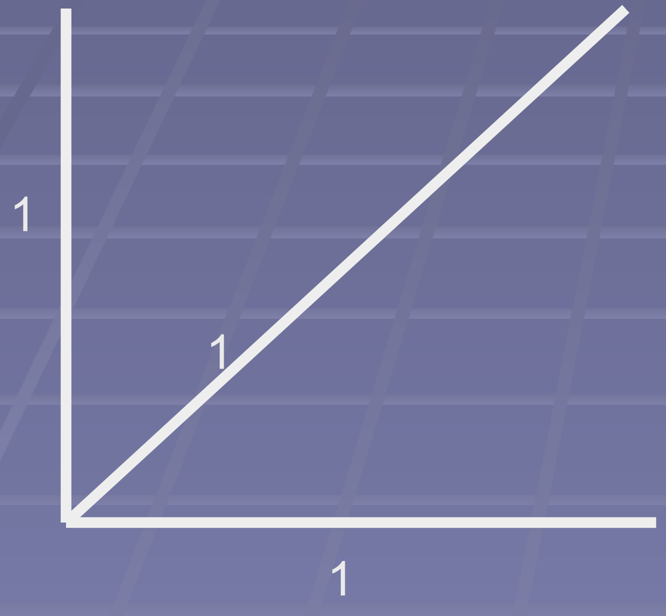




Un arbre T recouvrant de coût minimal 3.

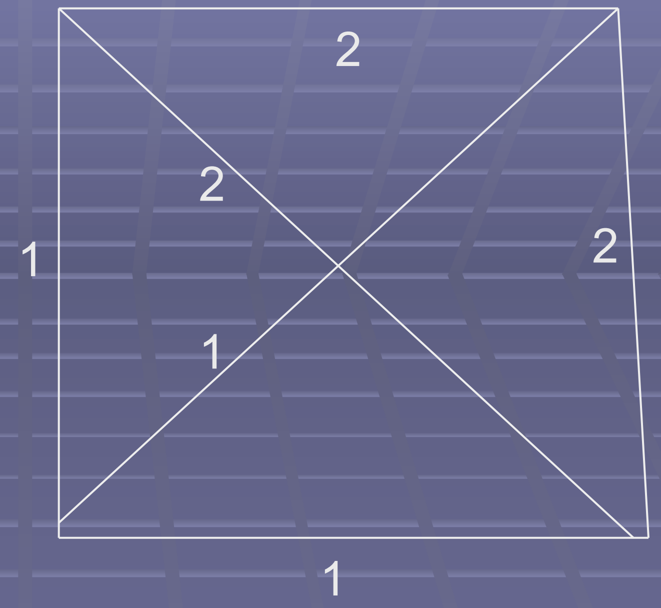
- **<u>Exemple</u>** ; graphe G :

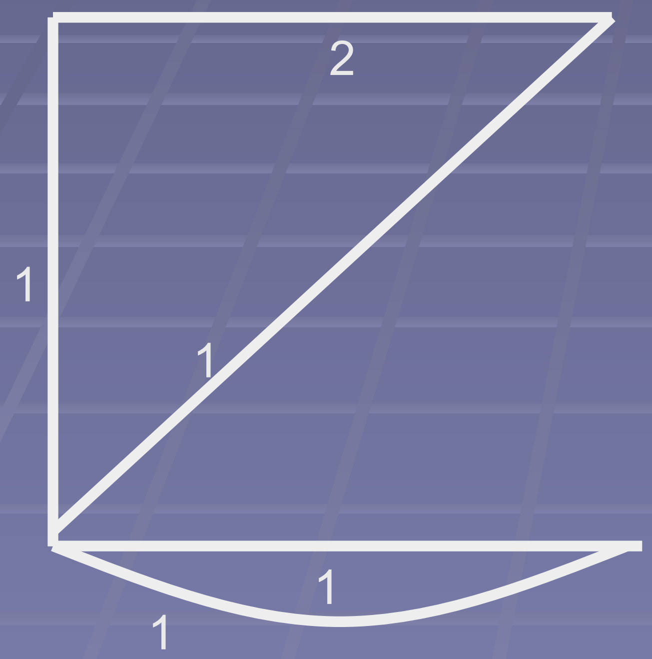




Un couplage parfait de coût minimal des sommets impairs de T.

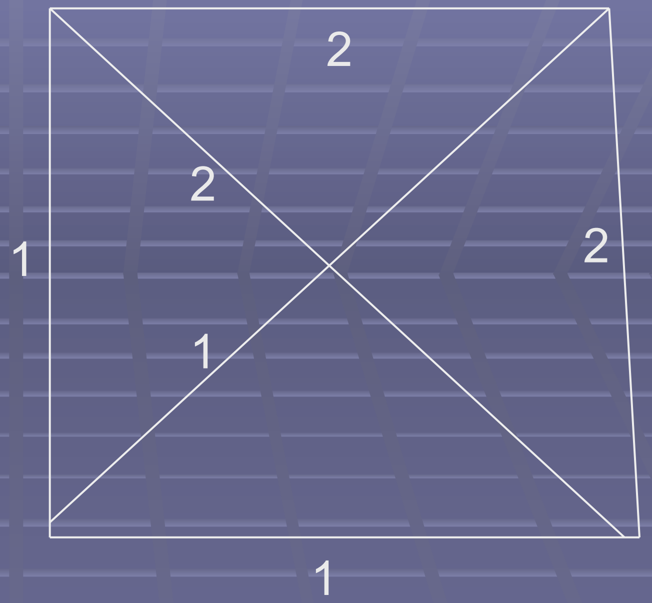
- **Exemple** ; graphe G :

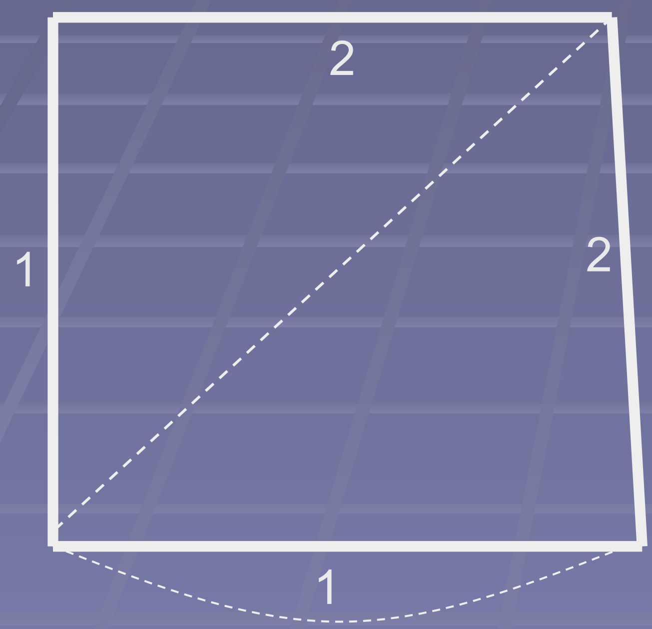




Un surgraphe de T contenant un cycle eulérien de coût minimal 6. Il contient un sommet de degré >2.

- **Exemple** ; graphe G :





Tous les sommets sont de degré 2 après « shortcut »: on obtient un cycle hamiltonien de coût 6.

- **Théorème :** Pour le Δ-PVC, la PRP de l'heuristique de Christofide est ≤ 1.5.

  Sa complexité est en $O(n^3)$.

Preuve : Soient O une solution au PVC, et H le cycle obtenu.

L'inégalité triangulaire $\Rightarrow c(H) \leq c(T)+c(C)$.

Or $c(T) \leq c(O)$ (car supprimer une arête de O donne un arbre recouvrant). Il suffit donc de montrer que :

$$c(C) \leq 0.5\, c(O)$$

Soit un circuit U dans G reliant les sommets I impairs de T, par des arêtes et dans l'ordre où ils apparaissent dans O.

$$U=(i_1,i_2)\,(i_2,i_3)\ldots(i_{p-1},i_p)$$

L'inégalité triangulaire $\Rightarrow$ chaque arête est géodésique

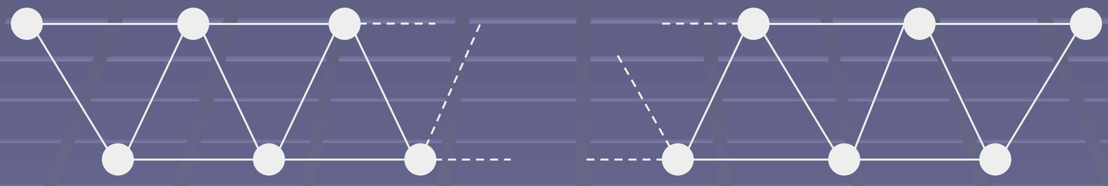
$\Rightarrow c(U) \leq c(O)$. Or U donne deux couplage parfaits de U : (i_1,i_2) ; (i_3,i_4) ; ... et (i_2,i_3) ; (i_4,i_5) ; ...

Donc l'un des deux est de coût au plus : 0.5 c(0).

Donc $c(C) \leq 0.5\, c(O)$ puisque c'est un couplage optimal. **cqfd**

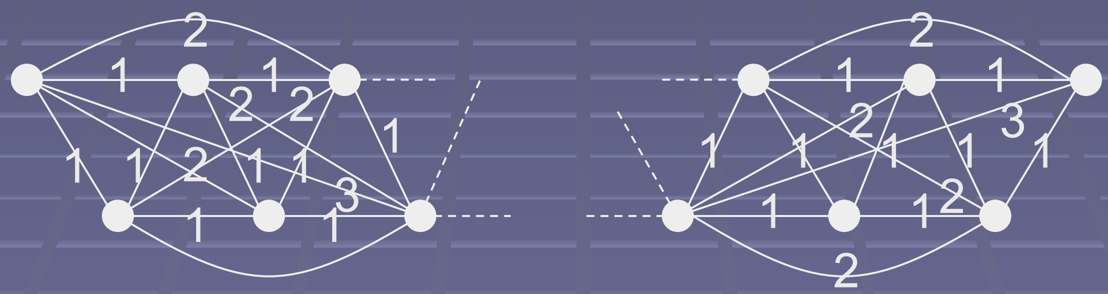
PRP(Christofide) = 1,5

- Considérer un graphe complet à n sommets dont les distances sont celles dans le graphe suivant :



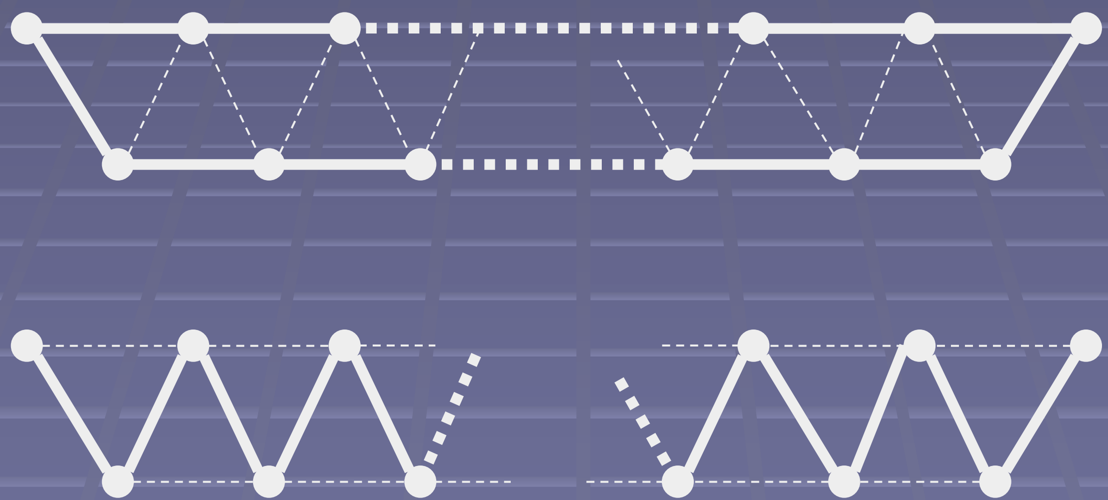
PRP(Christofide) = 1,5

- Considérer un graphe complet à n sommets dont les distances sont celles dans le graphe suivant :



PRP(Christofide) = 1,5

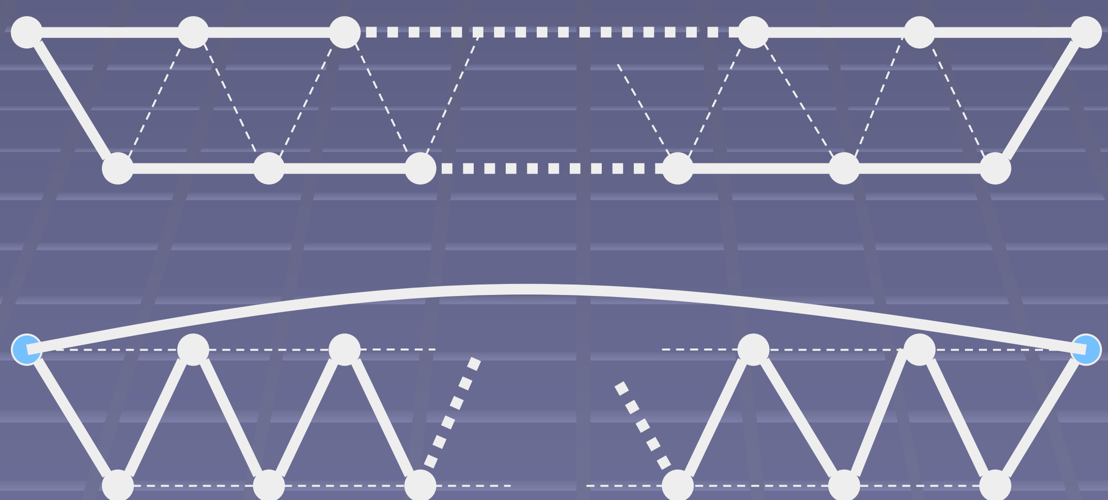
- Considérer un graphe complet à n sommets dont les distances sont celles dans le graphe suivant :



- Un cycle hamiltonien optimal a pour coût = n.
- Un arbre recouvrant de coût minimal = n-1.

PRP(Christofide) = 1,5

- Considérer un graphe complet à n sommets dont les distances sont celles dans le graphe suivant :



- Un cycle hamiltonien optimal a pour coût = n.
- Un arbre recouvrant de coût minimal = n-1.
- Christofide construit un cycle de coût n-1+(n-1)/2.
- Soit une $R_C=(3n-3)/2n \xrightarrow[n\to+\infty]{\nearrow} (3/2)^- \Rightarrow P_C=1,5$.