Circulations avec demandes et bornes inférieures

CM nº2 — Mobilité (M2 IMPAIRS)

Matěj Stehlík

19/1/2024

Et si on avait plusieurs sources et/ou puits ?

- Un aspect important du problème de flot max est qu'il n'y qu'une seule source s et un seul puits t.
- Supposons maintenant qu'il puisse y avoir un ensemble S de sources générant un flot, et un ensemble T de puits pouvant absorber le flot.
- Comme précédemment, nous supposons qu'il y a une capacité entière sur chaque arc.
- Avec des sources et des puits multiples, il est un peu difficile de décider quelle source ou quel puits privilégier dans un problème de maximisation.

Circulations admissibles

- Ainsi, au lieu de maximiser la valeur du flot, nous allons considérer un problème où :
  - les sources ont des valeurs d'offre fixes
  - les puits des valeurs de demande fixes
  - notre objectif est d'expédier le flot des sommets avec l'offre disponible vers ceux avec des demandes données.
- Imaginons, par exemple, que le réseau représente un système d'autoroutes ou de lignes de chemin de fer dans lequel nous voulons expédier des produits depuis des usines (qui ont une *offre*) vers des points de vente (qui ont une *demande*).
- Dans ce type de problème, nous ne cherchons pas à *maximiser* une valeur particulière ; nous voulons simplement *satisfaire* toute la demande en utilisant l'offre disponible.

Demandes et offres

- Soit $G = (V, E)$ un réseau avec des capacités sur les arcs.
- À chaque sommet $v \in V$ est associée une *demande* d_v.
- Si $d_v > 0$, cela indique que v a une demande de d_v ; le sommet est un puits et il souhaite recevoir d_v unités de flot en plus de celui qu'il n'en envoie.
- Si $d_v < 0$, cela indique que v a une offre de $-d_v$; le sommet est une source, et il souhaite envoyer $-d_v$ unités de flot de plus qu'il n'en reçoit.
- Si $d_v = 0$, alors le sommet v n'est ni une source ni un puits.
- Nous supposerons que toutes les capacités et demandes sont des entiers.

Contraintes de capacité et de demande

- Soit S l'ensemble de tous les sommets ayant une demande *négative* et T l'ensemble de tous les sommets ayant une demande *positive*.
- Une *circulation avec des demandes* $\{d_v\}$ est une fonction f qui attribue un nombre réel non négatif à chaque arc et qui satisfait les deux contraintes suivantes :

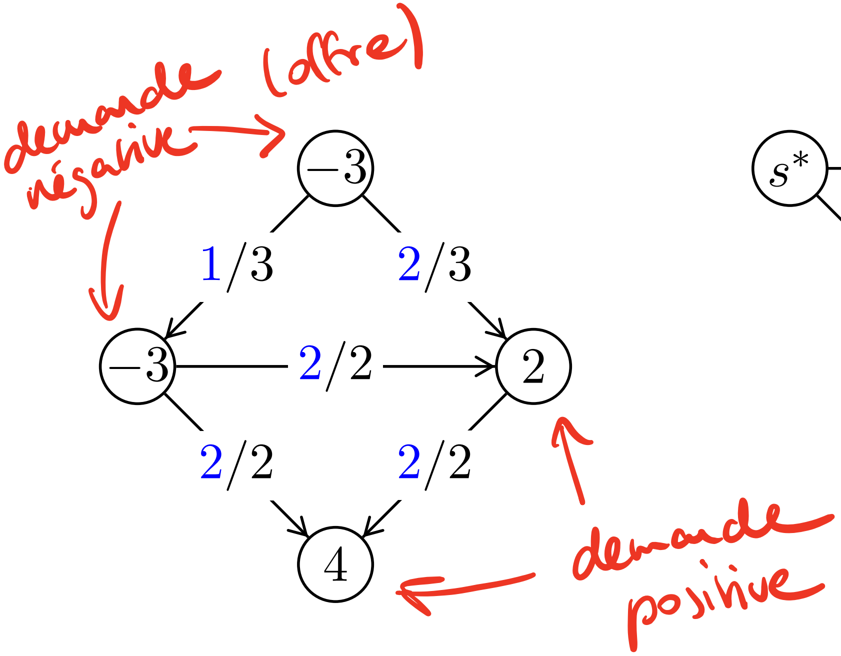
  Contraintes de capacité Pour chaque $e \in E$, on a $0 \leq f(e) \leq c_e$.

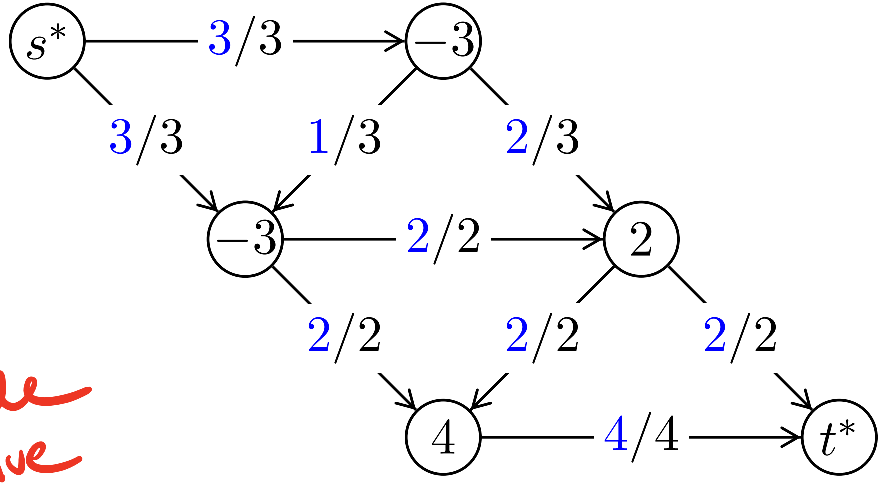
  Contraintes de demande Pour chaque $v \in V$, on a

$$\sum_{e \in \delta^-(v)} f(e) - \sum_{e \in \delta^+(v)} f(e) = d_v.$$

- Au lieu de considérer un problème de maximisation, nous nous intéressons à un problème de faisabilité : nous voulons savoir s'il existe une circulation qui satisfait les contraintes.

Exemple


demande négative (offre)
−3
1/3
2/3
−3
2/2
2
2/2
2/2
4
demande positive


s*
3/3
−3
3/3
1/3
2/3
−3
2/2
2
2/2
2/2
2/2
4
4/4
t*

Une condition nécessaire

Observation

admissible

S'il existe une circulation ~~réalisable~~ avec des demandes $\{d_v\}$, alors $\sum_{v \in V} d_v = 0$.

Démonstration

- Supposons qu'il existe une circulation f admissible.
- Alors $\sum_{v \in V} d_v = \sum_{v \in V} \left(f^{\text{in}}(v) - f^{\text{out}}(v)\right)$.
- La valeur $f(e)$ de chaque arc $e = (u, v)$ est comptée exactement deux fois : une fois dans $f^{\text{out}}(u)$ et une fois dans $f^{\text{in}}(v)$.
- Ces deux termes s'annulent, et puisque ceci est valable pour toutes les valeurs $f(e)$, la somme vaut 0.

Réduction au problème de flot max

- On peut réduire le problème de circulation admissible avec demandes $\{d_v\}$ au problème de flot max.
- On attache une "super source" s^* à chaque sommet de S et un "super puits" t^* à chaque sommet de T.
- Plus précisément, nous créons un graphe G' à partir de G en ajoutant de nouveaux sommets s^* et t^* à G.
- Pour chaque sommet $v \in T$, nous ajoutons un arc (u, t^*) avec la capacité d_u.
- Pour chaque sommet $u \in S$, nous ajoutons un arc (s^*, u) avec la capacité $-d_u$.

Équivalence des deux problèmes (1/2)

- Dans le graphe G', nous chercherons un $s^* - t^*$ flot max .
- Intuitivement, s^* alimente toutes les sources avec leur flot supplémentaire, et t^* siphonne le flot supplémentaire des puits.
- Il ne peut pas y avoir de $s^* - t^*$ flot dans G' d'une valeur supérieure à $D := \sum_{v:d(v)>0} d(v) = -\sum_{v:d(v)<0} d(v)$, puisque la coupe $(\{s^*\}, V \setminus \{s^*\})$ est de capacité D.
- S'il y a une circulation f admissible avec des demandes $\{d_v\}$ dans G, alors en envoyant un flot de valeur $-d_v$ dans chaque arc (s^*, v), et un flot de valeur d_v dans chaque arc (v, t^*), on obtient un $s^* - t^*$ flot dans G' de valeur D, il s'agit donc d'un flot max.

Équivalence des deux problèmes (2/2)

- Inversement, supposons qu'il existe un $s^* - t^*$ flot (maximal) dans G' de valeur D.
- Chaque arc sortant de s^*, et chaque arc entrant dans t^*, est complètement saturé.
- Ainsi, si l'on supprime ces arcs, on obtient une circulation f dans G avec $f^{\text{in}}(v) - f^{\text{out}}(v) = d_v$ pour chaque sommet v.
- De plus, s'il existe un flot de valeur D dans G', alors il existe un tel flot entier.

Théorème

Il existe une circulation admissible avec des demandes $\{d_v\}$ dans G ssi le $s^* - t^*$ flot maximal dans G' est de valeur D. Si toutes les capacités et demandes dans G sont entières et qu'il existe une circulation admissible, alors il existe une circulation admissible entière.

Caractérisation de graphes ayant une circulation admissible

- Dans le contexte des problèmes de circulation avec demandes, une coupe (A, B) est toute partition de l'ensemble de sommets V en deux ensembles, sans restriction sur le côté de la partition où se trouvent les sources et les puits.

Théorème

Il existe une circulation admissible avec demandes $\{d_v\}$ ssi

$$\sum_{v \in B} d_v \leq \text{cap}(A, B)$$

pour chaque coupe (A, B).

Circulations avec demandes et bornes inférieures

- Dans de nombreuses applications, en plus de satisfaire les demandes des différents sommets, on veut aussi forcer le flot à utiliser certains arcs.
- On peut placer des bornes inférieures sur les arcs, ainsi que les bornes supérieures habituelles imposées par les capacités des arcs.
- Considérons un réseau $G = (V, E)$ avec une capacité c_e et une borne inférieure ℓ_e sur chaque arc e.
- Nous supposerons que $0 \leq \ell_e \leq c_e$ pour chaque e.
- Comme précédemment, chaque sommet v aura également une demande d_v, qui peut être soit positive soit négative.
- Nous supposerons que toutes les demandes, capacités et bornes inférieures sont des entiers.

Contraintes de capacité et de demande

- Une borne inférieure ℓ_e signifie que la valeur du flot dans e doit être au moins ℓ_e.
- Ainsi, une circulation dans le réseau doit satisfaire les deux contraintes suivantes.

  Contraintes de capacité Pour chaque $e \in E$, on a $\ell_e \leq f(e) \leq c_e$.

  Contraintes de demande Pour chaque $v \in V$, on a

  $$f^{\text{in}}(v) - f^{\text{out}}(v) = d_v.$$

- Comme précédemment, il faut décider s'il existe une circulation admissible.

Réduction aux circulations sans bornes inférieures (1/2)

- On va réduire ce problème au problème de circulation avec des demandes sans bornes inférieures.
- Il faut envoyer au moins ℓ_e unités de flot dans chaque arc e.
- On définit une circulation initiale f_0 par $f_0(e) = \ell_e$.
- f_0 satisfait toutes les ~~conditions~~ contraintes de capacité (bornes inférieures et supérieures), mais il ne satisfait probablement pas toutes les conditions de demande

$$f_0^{\text{in}}(v) - f_0^{\text{out}}(v) = \sum_{e \in \delta^-(v)} \ell_e - \sum_{e \in \delta^+(v)} \ell_e =: L_v.$$

Réduction aux circulations sans bornes inférieures (2/2)

- Si $L_v = d_v$, alors la condition de demande à v est satisfaite.
- Si ce n'est pas le cas, on va superposer une circulation f_1 sur f_0 qui éliminera le déséquilibre restant en v.
- Il faut donc que $f_1^{\text{in}}(v) - f_1^{\text{out}}(v) = d_v - L_v$.
- Ayant déjà envoyé ℓ_e unités de flot dans chaque arc e, il en reste $c_e - \ell_e$ unités.
- Ces considérations motivent directement la construction suivante.
  - Le graphe G' a les mêmes sommets et arcs que G, mais pas de bornes inférieures.
  - La capacité de l'arc e est $c_e - \ell_e$.
  - La demande du sommet v est $d_v - L_v$.

Équivalence des deux problèmes (1/2)

Théorème

Il existe une circulation admissible dans G ssi il existe une circulation admissible dans G'. Si toutes les demandes, capacités et bornes inférieures dans G sont des entiers, et qu'il existe une circulation admissible, alors il existe une circulation réalisable entière dans G.

Démonstration (1/2)

admissible

- Supposons d'abord qu'il existe une circulation f' dans G'.
- Définissons une circulation f dans G par $f(e) = f'(e) + \ell_e$.
- f satisfait les conditions de capacité dans G, et

$$f^{\text{in}}(v) - f^{\text{out}}(v) = \sum_{e \in \delta^-(v)} (\ell_e + f'(e)) - \sum_{e \in \delta^+(v)} (\ell_e + f'(e)) = L_v + (d_v - L_v) = d_v$$

Équivalence des deux problèmes (2/2)

Démonstration (2/2)

- Donc f satisfait aussi les conditions de demande dans G.
- Inversement, supposons qu'il existe une circulation f dans G.
- Définissons une circulation f' dans G' par $f'(e) = f(e) - \ell_e$.
- f' satisfait les conditions de capacité dans G', et

$$(f')^{\text{in}}(v) - (f')^{\text{out}}(v) = \sum_{e \in \delta^-(v)} (f(e) - \ell_e) - \sum_{e \in \delta^+(v)} (f(e) - \ell_e) = d_v - L_v$$

- Donc f' satisfait aussi les conditions de demande dans G'.

Compagnies aériennes

- Les problèmes informatiques auxquels sont confrontés les grands transporteurs aériens sont extrêmement complexes.
- Ils doivent produire des programmes pour des milliers de routes chaque jour qui sont efficaces en termes d'utilisation de l'équipement, d'affectation des équipages, de satisfaction des clients et d'une foule d'autres facteurs (tout cela face à des problèmes imprévisibles comme la météo et les pannes).
- Les compagnies aériennes sont parmi les plus grands utilisateurs de techniques algorithmiques puissantes.
- Nous allons discuter d'un problème simplifié qui capture, de manière très claire, certaines des questions d'allocation des ressources qui se posent dans un tel contexte.

Le problème

- Il faut créer un planning de vols pour une flotte d'avions.
- Une étude de marché a identifié un ensemble de m segments de vol particuliers qui seraient très lucratifs si vous pouviez les desservir.
- Le segment de vol j est spécifié par quatre paramètres :
  1. l'aéroport d'origine
  2. l'aéroport de destination
  3. l'heure de départ
  4. l'heure d'arrivée.

Exemple de segments de vol lucratifs

1. Boston (BOS ; départ 6h) – Washington DC (DCA : arrivée 7h)
2. Philadelphie (PHL ; départ 7h) – Pittsburgh (PIT ; arrivée 8h)
3. Washington DC (DCA ; départ 8h) – Los Angeles (LAX ; arrivée 11h)
4. Philadelphie (PHL ; départ 11h) – San Francisco (SFO ; arrivée 14h)
5. San Francisco (SFO ; départ 14h15) – Seattle (SEA ; arrivée 15h15)
6. Las Vegas (LAS ; départ 17h) – Seattle (SEA ; arrivée 18h)

Utiliser un seul avion pour plusieurs segments

Il est possible d'utiliser un seul avion pour un segment de vol i, puis plus tard pour un segment de vol j, à condition que :

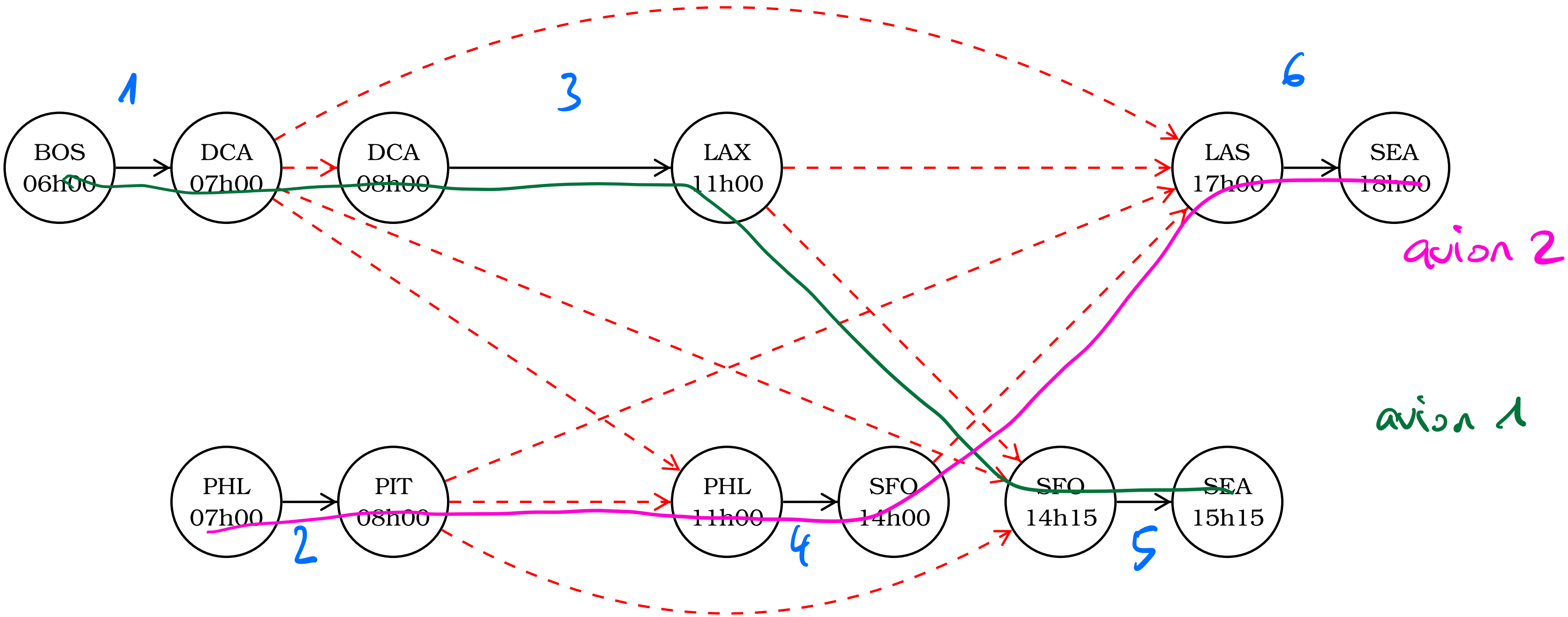
1. la destination de i est la même que l'origine de j, et qu'il y ait suffisamment de temps pour effectuer la maintenance de l'avion entre les deux ; ou bien
2. il est possible d'ajouter un segment de vol entre les deux, qui permette à l'avion d'aller de la destination de i à l'origine de j en laissant suffisamment de temps entre les deux.

Exemple

Par exemple, en supposant que le temps de maintenance est d'une heure, vous pouvez utiliser un seul avion pour les vols (1), (3) et (6) en insérant le vol suivant :

- Los Angeles (LAX ; départ 12h) – Las Vegas (LAS ; arrivée 13h)

Les vols lucratifs (en noir) et les vols accessibles (en rouge)



Formulation du problème (1/3)

- Nous dirons simplement que le vol j est *atteignable* depuis le vol i s'il est possible d'utiliser le même avion pour le vol i, puis plus tard pour le vol j également.
- Ainsi, selon les deux règles ci-dessus, nous pouvons facilement déterminer pour chaque paire (i, j) si le vol j est atteignable depuis le vol i.
- Il peut y avoir des règles plus complexes pour l'accessibilité : par exemple, la durée d'entretien nécessaire pourrait dépendre de l'aéroport, ou nous pourrions exiger que le segment de vol que vous insérez soit suffisamment rentable.

Formulation du problème (2/3)

- Néanmoins, nous pouvons traiter n'importe quel ensemble de règles avec notre définition : l'entrée du problème ne comprendra pas seulement les segments de vol, mais aussi une spécification des paires (i, j) pour lesquelles un vol ultérieur j est accessible à partir d'un vol antérieur i.
- Ces paires peuvent former un graphe acyclique orienté (DAG) arbitraire.
- L'objectif est de déterminer s'il est possible de desservir les m vols de votre liste initiale en utilisant au maximum k avions au total.
- Pour ce faire, vous devez trouver un moyen de réutiliser efficacement les avions pour de multiples vols.

Formulation du problème (3/3)

- Par exemple, revenons à l'exemple initial, et supposons que nous disposons de $k = 2$ avions.
- Si nous utilisons l'un des avions pour les vols (1), (3) et (6) comme proposé ci-dessus, nous ne pourrons pas desservir tous les vols (2), (4) et (5) avec l'autre avion, car il n'y aurait pas assez de temps de maintenance à San Francisco entre les vols (4) et (5).
- Cependant, il existe une solution différente.
- Un avion peut desservir les vols (1), (3) et (5) (en ajoutant un vol LAX-SFO), tandis que l'autre peut desservir les vols (2), (4) et (6) (en ajoutant les vols PIT-PHL et SFO-LAS).

Vers un algorithme efficace (1/2)

- Nous discutons maintenant d'un algorithme efficace qui peut résoudre des instances arbitraires du problème d'ordonnancement des compagnies aériennes, basé sur les flots dans un réseau.
- Nous verrons que les techniques de flots s'adaptent très naturellement à ce problème.
- La solution est basée sur l'idée suivante.
- Les unités de flot correspondront à des avions.
- Il y aura un arc pour chaque vol, et des limites de capacité supérieure et inférieure de 1 sur ces arcs pour exiger qu'exactement une unité de flot traverse cet arc.

Vers un algorithme efficace (2/2)

- En d'autres termes, chaque vol doit être desservi par l'un des avions.
- Si (u_i, v_i) est l'arc représentant le vol i, et (u_j, v_j) est l'arc représentant le vol j, et que le vol j est atteignable depuis le vol i, alors il y aura un arc de v_i à u_j avec une capacité de 1.
- De cette façon, une unité de flot peut traverser (u_i, v_i) et ensuite se déplacer directement vers (u_j, v_j).
- Nous l'étendons à un réseau en incluant une source et un puits.
- Nous allons maintenant décrire la construction en détail.

Réduction aux flots (1/2)

L'ensemble des sommets du graphe sous-jacent G est défini comme suit :

- Pour chaque vol i, le graphe G comportera les deux sommets u_i et v_i.
- G aura également un sommet source s et un sommet puits t.

L'ensemble des arcs de G est défini comme suit :

- Pour chaque i, il existe un arc (u_i, v_i) avec une borne inférieure de 1 et une capacité de 1. *(Chaque vol de la liste est desservi.)*
- Pour chaque i et j t.q. le vol j est atteignable à partir du vol i, il existe un arc (v_i, u_j) avec borne inférieure 0 et capacité 1. *(Le même avion peut effectuer les vols i et j.)*

Réduction aux flots (2/2)

- Pour chaque i, il y a un arc (s, u_i) avec borne inférieure 0 et capacité 1. *(Tout avion peut commencer la journée par le vol i.)*
- Pour chaque j, il y a un arc (v_j, t) avec borne inférieure 0 et capacité 1. *(Tout avion peut terminer la journée avec le vol j.)*
- Il y a un arc (s, t) avec borne inférieure 0 et capacité k. *(On n'est pas obligé d'utiliser tous les avions.)*

Finalement, les demandes sont défini comme suit :

- Le sommet s aura une demande de $-k$
- Le sommet t aura une demande de k.
- Tous les autres sommets auront une demande de 0.

Notre algorithme consiste à construire le réseau G et à rechercher une circulation admissible dans G.

Caractérisation de planning admissible (1/4)

Théorème

Les vols prioritaires peuvent être desservis en utilisant au plus k avions si et seulement si il existe une circulation admissible dans G.

Démonstration (1/4)

- Tout d'abord, supposons que tous les vols prioritaires peuvent être desservis en utilisant $k' \leq k$ avions.
- L'ensemble des vols effectués par chaque avion individuel définit un chemin P dans le réseau G, et nous envoyons une unité de flot sur chacun de ces chemins P.

Caractérisation de planning admissible (2/4)

Démonstration (2/4)

- Pour satisfaire la totalité des demandes en s et t, nous envoyons $k - k'$ unités de flot sur l'arc (s, t).
- La circulation résultante satisfait toutes les conditions de demande, de capacité et de borne inférieure.
- Inversement, considérons une circulation admissible dans le réseau G.
- Par la caractérisation de circulations admissibles, il existe une circulation admissible avec des valeurs de flot entières.
- Supposons que k' unités de flot soient envoyées sur des arcs autres que (s, t).

Caractérisation de planning admissible (3/4)

Démonstration (3/4)

- Puisque tous ces arcs ont capacité 1, et que la circulation est entière, chacun des arcs qui transporte du flot transporte exactement une unité de flot.
- Nous convertissons maintenant ceci en un planning.
- Considérons un arc (s, u_i) qui transporte une unité de flot.
- Par la conservation de flot, (u_i, v_i) transporte une unité de flot, et il existe un unique arc sortant de v_i qui transporte une unité de flot.
- Si nous continuons de cette manière, nous construisons un chemin P de s à t, de sorte que chaque ~~arête~~ arc de ce chemin transporte une unité de flot.

Caractérisation de planning admissible (4/4)

Démonstration (4/4)

- Nous pouvons appliquer cette construction à chaque arête de la forme (s, u_j) transportant une unité de flot ; de cette façon, nous produisons k' chemins de s à t, chacun étant constitué d'arcs qui transportent une unité de flot.
- Nous pouvons affecter un seul avion à chaque chemin P que nous créons de cette manière, pour effectuer tous les vols contenus dans ce chemin.

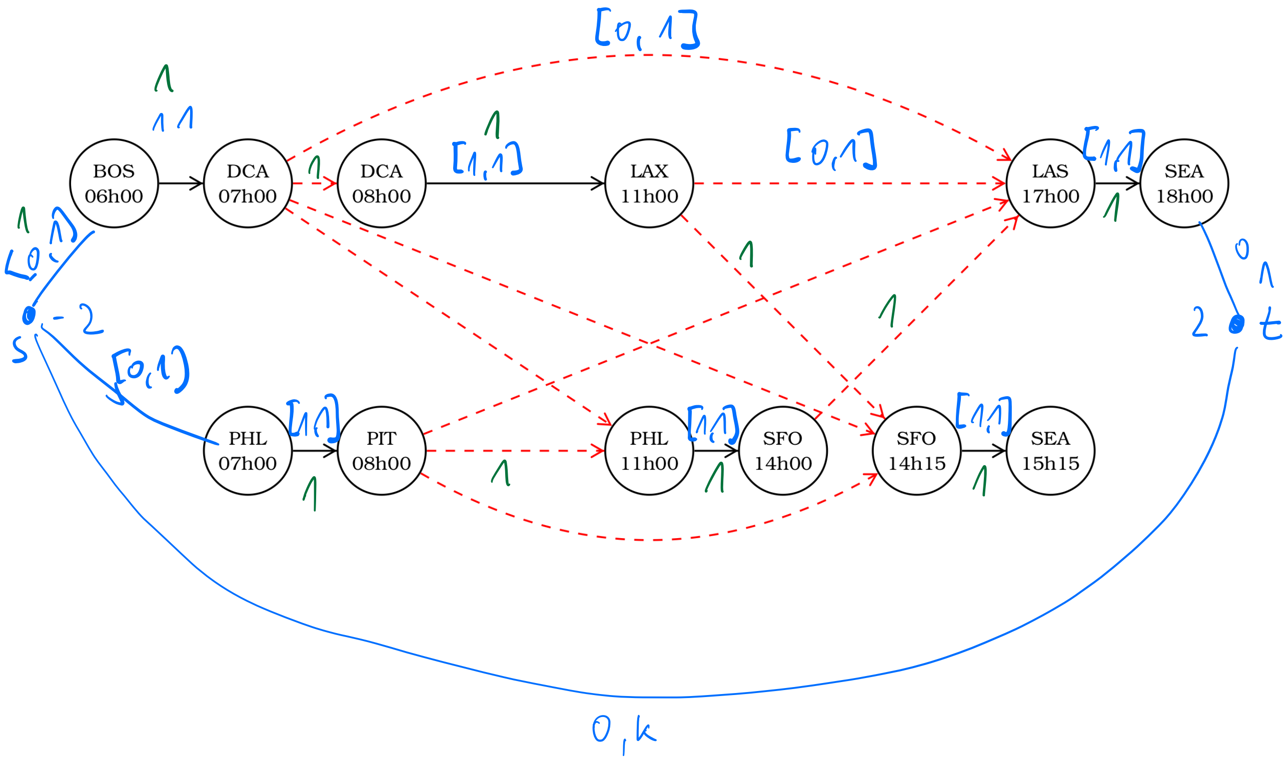
Aspects du problème que nous avons ignoré (1/2)

Notre problème simplifié ignore plusieurs aspects importants :

- Il ne tient pas compte du fait qu'un avion donné ne peut voler qu'un certain nombre d'heures avant de devoir être temporairement mis hors service pour une maintenance plus importante.
- Nous établissons un planning optimal pour un seul jour comme s'il n'y avait ni hier ni demain ; en fait, nous avons également besoin que les avions soient positionnés de manière optimale pour le début du jour $N+1$ à la fin du jour N.

Aspects du problème que nous avons ignoré (2/2)

- Les avions sont dotés d'équipages (êtres humains), ce qui introduit de nouvelles contraintes.
- La desserte d'un segment de vol particulier n'est pas une contrainte stricte ; l'objectif réel est plutôt d'optimiser les recettes, et nous pouvons donc choisir parmi de nombreux vols possibles à inclure dans notre planning (sans parler de la conception d'une bonne structure tarifaire pour les passagers) afin d'atteindre cet objectif.



BOS
06h00
DCA
07h00
DCA
08h00
LAX
11h00
LAS
17h00
SEA
18h00
PHL
07h00
PIT
08h00
PHL
11h00
SFO
14h00
SFO
14h15
SEA
15h15
s
-2
t
2
[0,1]
[1,1]
0,k