

Algorithmes de calcul des ordres d'appel

Hugo Gimbert

June 6, 2019

1 Cas d'un seul taux boursier

Soit C l'ensemble des candidats partitionné en deux sous-ensemble B les candidats boursiers et NB les candidats non-boursiers.

L'algorithme du calcul de l'ordre d'appel avec un seul taux prend en entrée un taux minimum de candidats boursiers $q_B \in [0, 1]$ et un classement pédagogique $\text{rang} : C \rightarrow 1 \dots |C|$ des candidats. On suppose que les rangs dans le classement pédagogique sont des entiers consécutifs à partir de 1 sans ex-aequo, c'est-à-dire que $\text{rang}(C) = \{1 \dots |C|\}$. Ce classement induit une permutation $(c_i)_{i \in 1 \dots |C|}$, c'est l'unique permutation telle que $\forall i \in 1 \dots |C|, \text{rang}(c_i) = i$.

1.1 Spécification

L'algorithme calcule une autre permutation $(d_k)_{k \in 1 \dots |C|}$ des candidats, appelée l'ordre d'appel, qui garantit les propriétés suivantes.

- (P1) Pour tout k , l'ensemble de candidats d_1, \dots, d_k contient au moins $q_B * k$ candidats boursiers ; ou sinon, aucun candidat parmi $d_{k+1} \dots d_{|C|}$ n'est boursier.
- (P2) Un candidat boursier qui a le rang r dans le classement pédagogique n'est jamais doublé par personne et aura donc un rang inférieur ou égal à r dans l'ordre d'appel.
- (P3) Un candidat non boursier qui a le rang r dans le classement pédagogique ne double jamais personne et aura un rang compris entre r et l'arrondi à l'entier supérieur de $r(1 + q_B/(1 - q_B))$ dans l'ordre d'appel.
- (P4) Comparé au classement pédagogique, l'ordre d'appel minimise le nombre d'inversions (distance de Kendall-tau), parmi ceux qui garantissent la première propriété.
- (P5) Si l'on munit l'ensemble des sélections ordonnées de candidats de l'ordre lexicographique induit par les classements, alors l'ordre d'appel est le maximum parmi toutes les sélections qui garantissent la première propriété.

1.2 Implémentation

L'algorithme commence par extraire de $(c_i)_{i \in 1 \dots |C|}$ les permutations correspondantes $(b_i)_{i \in 1 \dots |B|}$ des boursiers et $(nb_i)_{i \in 1 \dots |\text{NB}|}$ des non-boursiers. Elles sont caractérisées par

- $(b_i)_{i \in 1 \dots |B|}$ est une permutation de B ,
- $\forall 1 \leq i \leq j \leq |B|, \text{rang}(b_i) \leq \text{rang}(b_j)$.
- $(nb_i)_{i \in 1 \dots |\text{NB}|}$ est une permutation de NB ,
- $\forall 1 \leq i \leq j \leq |\text{NB}|, \text{rang}(nb_i) \leq \text{rang}(nb_j)$.

Le calcul est effectué en $n + 1$ étapes $0, 1, \dots, n$. Au début de l'étape k , la permutation $d_1 \dots d_k$ est déjà calculée, ainsi que deux indices $r_k(B) \in 0 \dots |B|$ et $r_k(\text{NB}) \in 0 \dots |\text{NB}|$. Initialement $k = 0$ et $r_0(B) = 0$ et $r_0(\text{NB}) = 0$. Le calcul se termine au début de l'étape $k = n$.

Les deux indices $r_k(B)$ et $r_k(\text{NB})$ sont des pointeurs dans les permutations $(b_i)_{i \in 1 \dots |B|}$ et $(nb_i)_{i \in 1 \dots |\text{NB}|}$ qui sont mis-à-jour de manière à garantir les invariants

$$\begin{aligned} B \cap \{d_1, \dots, d_k\} &= \{b_1, \dots, b_{r_k(B)}\} \\ \text{NB} \cap \{d_1, \dots, d_k\} &= \{nb_1, \dots, nb_{r_k(\text{NB})}\} . \end{aligned}$$

Le choix de d_{k+1} est fait à l'étape k comme suit.

On dit que le taux effectif de boursier est *contraignant* après $d_1 \dots d_k$, si ajouter un non-boursier à la permutation ferait passer le taux de boursiers en dessous du minimum fixé alors qu'il reste un boursier à sélectionner c'est-à-dire si

$$(r_k(B) < |B|) \text{ et } (r_k(B) < q_B * (k + 1)) .$$

Le choix de d_{k+1} est fait parmi un ou deux candidats *éligibles*. L'ensemble des candidats éligibles est calculé comme suit. Si il reste un boursier non-sélectionné, i.e. si $r_k(B) < |B|$, alors le candidat $b_{r_k(B)+1}$ est éligible. De plus, si le taux n'est pas contraignant après $d_1 \dots d_k$ et qu'il reste un non-boursier non-sélectionné, i.e. si $r_k(\text{NB}) < |\text{NB}|$, alors le candidat $nb_{r_k(\text{NB})+1}$ est éligible.

On choisit pour d_{k+1} le candidat le mieux classé pédagogiquement parmi les candidats éligibles.

- Si c'est le boursier $d_{k+1} = b_{r_k(B)+1}$ alors $r_{k+1}(B) = r_k(B) + 1$ et $r_{k+1}(\text{NB}) = r_k(\text{NB})$.
- Si c'est le non-boursier $d_{k+1} = nb_{r_k(\text{NB})+1}$ alors $r_{k+1}(B) = r_k(B)$ et $r_{k+1}(\text{NB}) = r_k(\text{NB}) + 1$.

1.3 Preuve des invariants

Pour démontrer les propriétés (P1)-(P5), il suffit de montrer que l'algorithme maintient les invariants suivants.

Pour tout $k \in 0 \dots |C|$ on note

$$g(k) = |B \cap \{d_1, \dots, d_k\}| \quad .$$

Donc le taux est contraignant après d_1, \dots, d_k si

$$(g(k) < |B|) \text{ et } (r_k(B) < q_B * (k + 1)) \quad .$$

L'algorithme satisfait les invariants suivants.

- (I0) $\forall 1 \leq i \leq k, g(i) = r_i(B)$.
- (I0') $\forall 1 \leq i \leq k$, si le taux n'est pas contraignant après d_1, \dots, d_{i-1} alors d_i est le candidat le mieux classé parmi $C \setminus \{d_1 \dots d_{i-1}\}$. Si le taux est contraignant après d_1, \dots, d_{i-1} alors d_i est le candidat le mieux classé parmi $B \setminus \{b_1 \dots b_{g(i-1)}\}$
- (I1) $d_1 \dots d_k$ est une permutation de $\{b_1 \dots b_{r_k(B)}\} \cup \{nb_1 \dots nb_{r_k(\text{NB})}\}$,
- (I2) $\forall 1 \leq i \leq k$, si d_i et d_k sont de même type alors $\text{rang}(d_i) \leq \text{rang}(d_k)$.
- (I3) $\forall 1 \leq i \leq k$, si d_k est boursier alors $\text{rang}(d_i) \leq \text{rang}(d_k) \leq k$.
- (I4) $\forall 1 \leq i \leq k$, si d_k est non-boursier et $\text{rang}(d_i) > \text{rang}(d_k)$ alors d_i est boursier et le taux est contraignant après d_1, \dots, d_{i-1} .
- (I5) $\forall 1 \leq i \leq k$, si d_i est non-boursier et $x_i = |\{j \in 1 \dots i \mid \text{rang}(d_j) > \text{rang}(d_i)\}|$ alors $x_i < q_B * i + 1$.
- (I6) $\forall 1 \leq i \leq k$, si le taux est contraignant après d_1, \dots, d_i alors $i < |C|$ et $d_{i+1} \in B$.
- (I7) $\forall 1 \leq i \leq k$, si $g(i) < q_B * i$ alors $g(i) = |B|$.

1.4 Preuve des propriétés à partir des invariants

Propriété (P1). Soit $k \in 1 \dots |C|$ tel que $g(k) < q_B * k$. Alors $g(k) = |B|$ d'après (I7). Donc d'après (I1),

$$\forall k \leq i \leq |C|, B = \{d_1 \dots d_i\} \cap B \quad .$$

donc par comptage, et puisque $(d_k)_{k \in 1 \dots |C|}$ est une permutation,

$$\forall k < i \leq |C|, d_i \notin B \quad .$$

Propriété (P2). Immédiate par (I3).

Propriété (P3). D'après (I5) $x_k < q_B * k + 1$. D'après (I2) et (I3), $1 \dots \text{rang}(d_k) \subseteq \text{rang}(\{d_1 \dots d_k\})$. Donc $k = \text{rang}(d_k) + x_k$. Finalement $k < \text{rang}(d_k) + q_B * k + 1$ donc $\text{rang}(d_k) \leq k < 1 + \text{rang}(d_k)/(1 - q_B)$, i.e. (P3).

Propriété (P5). Soit $(d'_k)_{k \in 1 \dots |C|}$ une autre permutation de C différente de $(d_k)_{k \in 1 \dots |C|}$ et qui respecte la propriété (P1). Soit k le plus petit indice où les deux permutations diffèrent. On veut montrer que $\text{rang}(d_k) < \text{rang}(d'_k)$.

Il y a deux cas, selon que le taux est contraignant ou non après d_1, \dots, d_{k-1} .

Si le taux n'est pas contraignant après d_1, \dots, d_{k-1} , alors d'après (I0'), d_k est le candidat le mieux classé parmi $C \setminus \{d_1 \dots d_{k-1}\}$ et puisque $\{d_1 \dots d_{k-1}\} = \{d'_1 \dots d'_{k-1}\}$ alors $d'_k \in C \setminus \{d_1 \dots d_{k-1}\}$ donc en particulier d_k est mieux classé que d'_k i.e. $\text{rang}(d_k) < \text{rang}(d'_k)$.

Si le taux est contraignant après d_1, \dots, d_{k-1} alors d'après (I0'), d_k est le candidat le mieux classé parmi $B \setminus \{b_1 \dots b_{g(k-1)}\}$. De plus d'_k est nécessairement boursier car le taux est contraignant et $(d'_k)_{k \in 1 \dots |C|}$ respecte (P1) donc puisque $\{d_1 \dots d_{k-1}\} = \{d'_1 \dots d'_{k-1}\}$ alors $d'_k \in B \setminus \{b_1 \dots b_{g(k-1)}\}$ et en particulier d_k est mieux classé que d'_k i.e. $\text{rang}(d_k) < \text{rang}(d'_k)$.

Propriété (P4). Soit $(d'_k)_{k \in 1 \dots |C|}$ qui minimise le nombre d'inversions

$$I' = |\{i \in 1 \dots |C|, j \in 1 \dots |C|, (i < j \wedge \text{rang}(d'_i) > \text{rang}(d'_j)) \vee (i > j \wedge \text{rang}(d'_i) < \text{rang}(d'_j))\}|$$

parmi toutes les permutations qui respectent (P1).

On veut montrer que $(d_k)_{k \in 1 \dots |C|} = (d'_k)_{k \in 1 \dots |C|}$. A contrario, supposons que ce soit faux, et soit k le plus petit indice où les deux permutations diffèrent et $d = d_k$ et $d' = d'_k$.

Puisque échanger deux boursiers ou deux non-boursiers préserve (P1) alors $(d'_k)_{k \in 1 \dots |C|}$ vérifie la propriété (I2) avec $k = |C|$.

Puisque échanger un boursier et un non-boursier en faisant remonter le boursier préserve (P1) alors $(d'_k)_{k \in 1 \dots |C|}$ vérifie la propriété (I3) avec $k = |C|$.

Il y a trois cas.

- Supposons d et d' de même type. D'après (I2), d est le mieux classé des candidats de sa catégorie dans $|C| \setminus \{d_1, \dots, d_{k-1}\}$. Et d' est le mieux classé des candidats de sa catégorie dans $|C| \setminus \{d'_1, \dots, d'_{k-1}\}$. On conclut puisque $\{d_1, \dots, d_{k-1}\} = \{d'_1, \dots, d'_{k-1}\}$.
- Supposons d est boursier. Remarquons que ni d ni d' n'apparaissent dans $d_1, \dots, d_{k-1} = d'_1, \dots, d'_{k-1}$. D'après (P5) $\text{rang}(d) < \text{rang}(d')$, donc on pourrait diminuer le nombre d'inversions dans $(d'_k)_{k \in 1 \dots |C|}$ en échangeant d et d' . Cela ferait remonter le boursier d et descendre le non-boursier d' donc n'affecterait pas (P1). Contradiction avec la minimalité de $(d'_k)_{k \in 1 \dots |C|}$.
- Reste le cas où d n'est pas boursier mais d' l'est. Alors le taux de boursier n'est pas contraignant après $d_1 \dots d_{k-1}$. Soit i le rang de d dans $(d'_k)_{k \in 1 \dots |C|}$. Puisque $d_1, \dots, d_{k-1} = d'_1, \dots, d'_{k-1}$ et $d = d_k$ et $d \neq d'$ alors

$i > k$. Puisque $(d'_k)_{k \in 1 \dots |C|}$ vérifie la propriété (I2) alors il n'y a que des boursiers dans la suite d'_k, \dots, d'_{i-1} . Que se passe t'il si on échange d et d' dans $(d'_k)_{k \in 1 \dots |C|}$? Déjà, cela diminue le nombre d'inversions: puisque $(d_k)_{k \in 1 \dots |C|}$ vérifie (I3) alors tous les boursiers dans d'_k, \dots, d'_{i-1} sont moins bien classés que d . De plus la nouvelle permutation respecté (P1) car $(d_k)_{k \in 1 \dots |C|}$ et $(d'_k)_{k \in 1 \dots |C|}$ respectent toutes deux (P1) (détails omis...).