Linear inequalities for neighborhood based dominance properties for the common due-date scheduling problem

Anne-Elisabeth Falq[1], Pierre Fouilhoux[1] and Safia Kedad-Sidhoum[2]

[1] Sorbonne Université, CNRS, LIP6, 4 place Jussieu, 75005 Paris, France
anne-elisabeth.falq@lip6.fr, pierre.fouilhoux@lip6.fr
[2] CNAM, CEDRIC, 292 rue Saint Martin, 75003 Paris, France safia.kedad_sidhoum@cnam.fr

Keywords: just-in-time scheduling, dominance properties, integer linear programming.

1 The common due-date problem

We consider a set of n tasks J that have to be processed non-preemptively on a single machine around a common due-date d. Given for each task $j \in J$ a processing time p_j and a unitary earliness (resp. tardiness) penalty α_j (resp. β_j), the problem denoted $1 \,|\,|\, \sum \alpha_j \,[d - C_j]^+ + \beta_j \,[C_j - d]^+$, aims at finding a feasible schedule that minimizes the sum of earliness-tardiness penalties.

When $d \geq \sum p_j$, the due date is said **unrestrictive**, and the problem is NP-hard, even if penalties are symmetric, *i.e.* $\alpha_j = \beta_j$ for all $j \in J$ (Hall and Posner 1991). In the general case, the problem is NP-hard, even if the task penalties are equal, *i.e.* $\alpha_j = \beta_j$ for all $j \in J$ (Hoogeveen and van de Velde 1991). In both cases to the dynamic programming algorithms proposed in (Hall and Posner 1991, Hoogeveen and van de Velde 1991). A heuristic method together with a benchmark is provided in (Biskup and Feldmann 2001). These instances are efficiently solved by an exact method proposed in (F. Sourd 2009).

In this work, we focus on the problem with an unrestrictive due-date. We propose a compact integer linear program modeling it. To improve the efficiency of this formulation, we propose a new type of linear inequalities translating some neighborhood based dominance properties. Moreover, for sake of brevity, we assume that the α-ratios α_j/p_j for $j \in J$ are different, as well as the β-ratios β_j/p_j. Nevertheless, the following results are still true without this assumption.

2 A compact linear formulation based on structural dominance properties

In a given schedule, a task is **early** (resp. **tardy**), if it completes before or at d (resp. after d), and a task is **on-time** if it completes exactly at time d. A schedule having an on-time task is said **V-shaped**, if early (resp. tardy) tasks are ordered by increasing α-ratios (resp. decreasing β-ratios). A schedule is called a **block** if it presents no idle time.

For the unrestrictive case, V-shaped blocks having an on-time task are dominant, which means that there exists an optimal solution within this set of schedules (Hall and Posner 1991). Using this dominance property, a schedule can be completely described by the partition between early and tardy tasks.
Indeed, if the set of early tasks E is given, the set of tardy tasks $T = J \setminus E$ is also fixed, and the earliness e_u (resp. the tardiness t_u) of any task $u \in J$, can be deduced as follows:

$$e_u = \begin{cases} p\big(A(u) \cap E\big) & \text{if } u \in E \\ 0 & \text{otherwise} \end{cases} \qquad t_u = \begin{cases} p\big(B(u) \cap T\big) & \text{if } u \in T \\ 0 & \text{otherwise} \end{cases}$$

where $p(S) = \sum_{j \in S} p_j$ for any $S \subseteq J$, $A(u) = \left\{ j \in J \mid \frac{\alpha_j}{p_j} > \frac{\alpha_u}{p_u} \right\}$ and $B(u) = \left\{ j \in J \mid \frac{\beta_j}{p_j} > \frac{\beta_u}{p_u} \right\}$.

Note that, for each task u, the sets $A(u)$ and $B(u)$ are defined from the instance, so they can be pre-computed. We introduce, for each $u \in J$, $\bar{A}(u) = J \setminus \big(A(u) \cup \{u\}\big)$ and $\bar{B}(u) = J \setminus \big(B(u) \cup \{u\}\big)$.

Let us consider a boolean variable δ_j for each $j \in J$ indicating if task j is early. *i.e.* a vector $\delta \in \{0,1\}^J$ encodes the partition $\big(E = \{j \in J \mid \delta_j = 1\},\ T = \{j \in J \mid \delta_j = 0\}\big)$. Although these variables are sufficient to encode solutions, additional boolean variables are introduced to replace quadratic terms appearing in the earliness and tardiness expression. Since these terms are only products of boolean variables, we use the classical linearization from (R. Fortet 1959) : for each couple in $J^< = \big\{(i,j) \in J^2 \mid i < j\big\}$, we add a new boolean variable $X_{i,j}$ and the four following inequalities coupling it with variables δ_i and δ_j.

$$\forall (i,j) \in J^<, \quad X_{i,j} \geqslant \delta_i - \delta_j \tag{1}$$
$$\forall (i,j) \in J^<, \quad X_{i,j} \geqslant \delta_j - \delta_i \tag{2}$$
$$\forall (i,j) \in J^<, \quad X_{i,j} \leqslant \delta_i + \delta_j \tag{3}$$
$$\forall (i,j) \in J^<, \quad X_{i,j} \leqslant 2 - (\delta_i + \delta_j) \tag{4}$$

If $(\delta, X) \in \{0,1\}^J \times \{0,1\}^{J^<}$ satisfies inequalities $(1)-(4)$, then $X_{i,j}$ indicates if $\delta_i \neq \delta_j$, and more importantly $\delta_i \delta_j = \delta_i + \delta_j - X_{i,j}$ and $(1-\delta_i)(1-\delta_j) = 2 - \delta_i - \delta_j - X_{i,j}$. The objective function reduces then to the following linear function:

$$f(\delta, X) = \sum_{u \in J} \alpha_u \left(\sum_{j \in A(u)} p_j \, \frac{\delta_j + \delta_u - X_{j,u}}{2} \right) + \beta_u \left((1-\delta_u)\, p_u + \sum_{j \in B(u)} p_j \, \frac{2 - \delta_j - \delta_u - X_{j,u}}{2} \right)$$

By introducing the polyhedron $P = \Big\{ (\delta, X) \in [0,1]^J \times [0,1]^{J^<} \mid (1)-(4) \Big\}$, and denoting $int(P)$ its integer points, the problem can be formulated as a linear integer program (A-E. Falq, P. Fouilhoux and S. Kedad-Sidhoum 2019):

$$(F) \quad \min_{(\delta, X) \in int(P)} f(\delta, X)$$

Since it has exactly $n + n(n-1)/2$ boolean variables and $4n(n-1)/2$ inequalities, (F) is a compact formulation. Note that no linear inequalities are needed to ensure the task non-overlapping since it is handled through the encoding.

3 Linear inequalities for neighborhood based dominance properties

It is common, in local search procedures, to slightly change a solution $\mathcal{S}$ to obtain a new one $\mathcal{S}'$, called a **neighbor** of $\mathcal{S}$. If the neighbor $\mathcal{S}'$ is better, (*i.e.* if it has a smaller total penalty in our case), we say that $\mathcal{S}$ is dominated (by $\mathcal{S}'$), it follows that $\mathcal{S}$ cannot be optimal.

This simple observation leads to a dominance property for any **neighborhood** $\mathcal{N}$ which associates to a solution the set of its neighbors. A solution $\mathcal{S}$ is said **$\mathcal{N}$-dominated** if there exists $\mathcal{S}' \in \mathcal{N}(\mathcal{S})$ which is strictly better than $\mathcal{S}$. Hence solutions which are not $\mathcal{N}$-dominated are dominant.

Here, as a schedule is encoded by a partition (E,T) between early and tardy tasks, we consider two operations providing a neighbor (E',T'):

- the **insertion** operation, which consists in inserting an early task on the tardy side *i.e.* $E' = E \setminus \{u\}$ and $T' = T \cup \{u\}$ for some $u \in E$, or conversely in inserting a tardy task on the early side *i.e.* $E' = E \cup \{u\}$ and $T' = T \setminus \{u\}$ for some $u \in T$,
- the **swap** operation, which consists in inserting an early task on the tardy side while a tardy task is inserted on the early side *i.e.* $E' = E \setminus \{u\} \cup \{v\}$ and $T' = T \setminus \{v\} \cup \{u\}$ for some $(u,v) \in E \times T$.

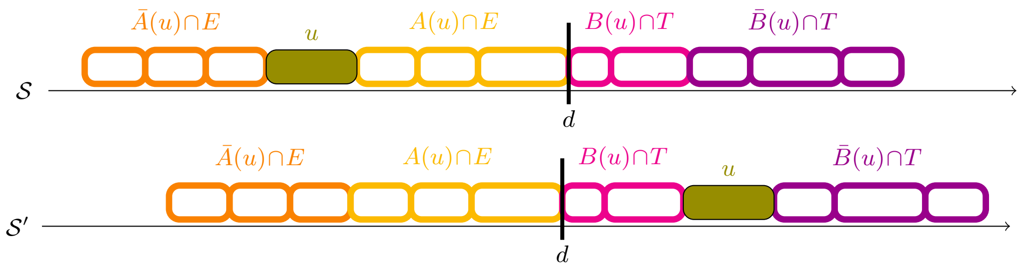

Fig. 1. Insertion of an early task u on the tardy side of a schedule

Figure 1 illustrates the insertion of an early task on the tardy side. Let us fix a task $u \in J$. The top part of the scheme shows the general form of an arbitrary schedule $\mathcal{S}$ in which u is an early task. The bottom part shows the general form of the neighbor $\mathcal{S}'$ of $\mathcal{S}$ obtained by inserting of u on the tardy side. Considering solutions as schedules (rather than partitions) allows to easily express the penalty variation between $\mathcal{S}$ and $\mathcal{S}'$ as follows:

$$-\alpha_u\, p\big(\underline{A(u)\cap E}\big) + \beta_u \Big(p\big(\underline{B(u)\cap T}\big) + p_u\Big) - p_u\, \alpha\big(\underline{\bar{A}(u)\cap E}\big) + p_u\, \beta\big(\underline{\bar{B}(u)\cap T}\big)$$

Using this the penalty variation expression, we produce a linear inequality

• which cuts **all** schedules in which u is early and which are dominated by the schedule obtained by inserting u on the tardy side,

• which is valid for any other schedule, in particular for all optimal schedules since they are non-dominated.

Two elements allow us to produce such an inequality. First, assuming that u is early in the schedule encoded by a vector (δ, X) is equivalent to assume that the linear term $1-\delta_u$ equal zero. Secondly, if (δ, X) encodes a schedule where u is early, the penalty variation induced by the insertion of u on the tardy side, denoted $\Delta_u^{av}(\delta)$ is linear in δ:

$$\Delta_u^{av}(\delta) = -\alpha_u \sum_{i\in A(u)} p_i\, \delta_i \;+\; \beta_u \sum_{i\in B(u)} p_i(1-\delta_i) + \beta_u p_u \;+\; p_u\Big(\sum_{i\in \bar{B}(u)} \beta_i\,(1-\delta_i) - \sum_{i\in \bar{A}(u)} \alpha_i\, \delta_i\Big)$$

and bounded by a constant:

$$\forall \delta \in \{0,1\}^J,\; -\Delta_u^{av}(\delta) \leqslant M_u^{av} \;\text{ where }\; M_u^{av} = \alpha_u\, p\big(A(u)\big) - \beta_u p_u + p_u\, \alpha\big(\bar{A}(u)\big)$$

We finally deduce the following inequality, which translates the dominance of the set of schedules non-dominated by the insertion of u:

$$\Delta_u^{av}(\delta) \geqslant -M_u^{av}\,(1-\delta_u) \;\; (5_u)$$

Following the same approach, we produce a similar inequality (6_u) cutting exactly the schedules in which u is tardy, and dominated by inserting u on the early side. We also produce an inequality $(7_{u,v})$, for given $v \neq u$, cutting exactly the schedules in which u is early, v is tardy, and dominated by swapping u and v.

Note that inequalities of family (5), (6) and (7) are not standard reinforcement inequalities. Classically, valid inequalities are added to cut extreme points which are not integer and then do not encode a feasible solution, since they correspond to a too optimistic value. On the contrary, these dominance inequalities cut some integer points which encode feasible solutions because they correspond to dominated, and then non-optimal, schedules.

4 Exact resolution and rounding heuristic

Let us introduce the polyhedron reinforced by the previous dominance inequalities $P' = \Big\{(\delta, X) \in [0,1]^J \times [0,1]^{J^<} \;\mid\; (1)-(4),\, \forall u \in J, (5_u), (6_u),\;\; \forall (u,v) \in J^<, (7_{u,v}), (7_{v,u})\Big\}$, and the associated formulation : $(F')\quad \min_{(\delta,X)\in int(P')} f(\delta, X)$.

Theoretically, we know that both formulations (F) and (F') give the same value.To compare them from a practical point of view, we implement them using a linear solver (CPLEX version 12.6.3.0), and test them on the benchmark proposed by (Biskup and Feldmann 2001). Under a time limit of one hour, formulation (F) using all CPLEX features allows to exactly solve instances up with [50] tasks, while formulation (F') without any CPLEX features allows to exactly solve instances up with [150] tasks.

Although designed for exact solving, (F) (resp. (F')) can be used to obtain a lower bound, by solving its linear relaxation denoted $\bar{F}$ (resp. $\bar{F}'$), and to obtain an upper bound together with a feasible schedule, by rounding the fractional solution of $\bar{F}$ (resp. $\bar{F}'$).

A first rounding procedure consists in rounding vector δ and then fixing X accordingly so that we obtain $\hat{x} \in int(P)$. We then obtain a feasible schedule and an upper bound UB1 for (F) (resp. UB1' for (F')). Note that $\hat{x}$ can violate some dominance inequalities, then $\hat{x} \notin P'$ (that implies in particular that CPLEX does not accept $\hat{x}$ as an incumbent solution). So we add a **repairing phase**, which consists in applying swap and insert operations as long as it is possible, *i.e.* while an insert inequality or a swap inequality is violated, meaning that an insert or a swap operation strictly improves the solution. We finally obtain a non-dominated schedule and a better upper bound UB2 (resp. UB2').

Note that this repairing phase can also be applied to the heuristic solution provided by the Biskup and Feldmann algorithm (Biskup and Feldmann 2001), which possibly transforms their upper bound UB3 in a better upper bound UB4. We compare experimentally these four upper bounds and show that UB2, UB2' and UB4 are very strong : they are exact for 45 over 50 instances (of size up to 150), and the average gap to the optimal value for the 5 other instances is less than 0,1%.

References

D. Biskup and M. Feldmann, 2001, "Benchmarks for scheduling on a single machine against restrictive and unrestrictive common due dates", *Computers and operations research*, 28:787–801

A-E. Falq, P. Fouilhoux and S. Kedad-Sidhoum, 2019, Mixed integer formulations using natural variables for single machine scheduling around a common due date. *CoRR*, abs/1901.06880, 2019.

R. Fortet, 1959, "L'Algèbre de Boole et ses applications en Recherche Opérationnelle", *Cahiers du Centre d'Études en Recherche Opérationnelle*, 4:5

N.G. Hall and M.E. Posner, 1991, "Earliness-tardiness scheduling problems, I: weighted deviation of completion times about a common due date", *Operations Research*, 39(5):836–846.

J.A. Hoogeveen and S.L. van de Velde, 1991, "Scheduling around a small common due date", *European Journal of Operational Research*, Vol 55:237–242,

F. Sourd, 2009, "New exact algorithms for one-machine earliness-tardiness scheduling", *INFORMS Journal on Computing*, 21(1):167–175