Norbert Ascheuer · Matteo Fischetti · Martin Grötschel

Solving the Asymmetric Travelling Salesman Problem with time windows by branch-and-cut

1. Introduction

This paper presents a computational study of the following time-constrained version of the asymmetric travelling salesman problem (ATSP): Consider a directed graph $D := (V \cup \{0\}, A)$ on $n + 1$ nodes. Node 0 is the starting node (depot) for a salesman. With each arc $(i, j) \in A$, an arc duration $c_{ij} > 0$ is associated. Furthermore, assume that for each node $i \in V$, a processing time $p_i \geq 0$, a release date $r_i \geq 0$, and a due date $d_i \geq r_i$ are given. The release date $r_i$ denotes the earliest possible (and the due date $d_i$ the latest possible) starting time for visiting (processing) node $i \in V$. For the depot node 0 we assume $r_0 = d_0 = 0$. The processing time $p_i$ represents the elapsed time between the arrival and the departure at node $i$. Throughout this paper we assume that arc durations, processing times, release dates, and due dates are nonnegative integer values. For due dates we also allow $d_i = \infty$. The interval $[r_i, d_i]$ is called the time window of node $i$, the width of the time window is given by $d_i - r_i$. The time window for node $i \in V$ is called active, if $r_i > 0$ or $d_i < \infty$. A time window $[0, \infty)$ is called relaxed. The problem is to find a sequence of the nodes (starting at the depot node 0 at time 0 and ending at node 0) with minimal cost such that for every node $i \in V$ the...
arrival time \( t_i \) at node \( i \in V \) lies within the given time window \([r_i, d_i]\). In our case waiting times are allowed, i.e., one may arrive at a node \( i \in V \) earlier than \( r_i \) and wait until the node is “released” at time \( r_i \). Waiting has no influence on the cost of a solution.

The ATSP-TW reduces to the ATSP if \( p_i = 0, r_i = 0 \), and \( d_i = +\infty \) for every \( i \in V \). Therefore, the ATSP-TW with general time windows is \( \mathcal{NP} \)–hard. Indeed, it is strongly \( \mathcal{NP} \)–complete to find a feasible solution for the ATSP-TW [38]. Furthermore, Tsitsiklis [43] showed that the symmetric version TSP-TW with general time windows is strongly \( \mathcal{NP} \)–complete, even if the underlying graph \( G \) is a path and all processing times equal 0.

It is apparent that the ATSP-TW can be formulated as a dynamic program and that it can be attacked by various branch–and–bound and other enumerative techniques of integer programming. We have chosen a polyhedral approach. In this case, the ATSP-TW is formulated as an integer linear program that is solved by a cutting plane algorithm. The cutting plane algorithm is based on cuts derived from the investigation of an associated polytope.

The reasons for this line of approach were threefold. First, the polyhedral approach to the symmetric TSP, see [1, 28, 36], and the ATSP, see [23], has been extremely successful and we hoped that this “positive behaviour” might also show up in our ATSP variant. Second, we could make use of separation algorithms that had been developed in related projects. Third, the application on which we focus on here, was part of a larger project where LP based techniques were the workhorse of the algorithmic approach.

This paper reports part of the results of an investigation for Siemens Nixdorf Informationssysteme (SNI) aiming at the (global) optimization of a semiautomatic PC manufacturing process, see Ascheuer [3] for details. One of the individual optimization problems arising here is the task to schedule the stacker cranes of automatic warehouses. In this case the stacker crane optimization can be formulated as an online ATSP, see Ascheuer et al. [5]. We have developed several online ATSP heuristics. To check the quality of the solutions of the online heuristics it is necessary to solve (offline) ATSPs with additional constraints, in particular, with time windows, see [6]. The instances from this stacker crane application, supplied by SNI, provide the basic data of the present computational study.

One problem with our line of attack was that there is no “natural” IP formulation of the ATSP-TW. There are many possibilities to model the ATSP-TW as an integer or mixed–integer linear program; and there is, at least to our knowledge, no way to tell which model holds the best computational perspectives. We have chosen to concentrate on three different models, explained in Sect. 3, and to let “computational experience” decide which to use in practice.

The paper is organized as follows. In Sect. 2 we briefly sketch related problems and work. In Sect. 3 we introduce the notation used throughout the paper and describe the three integer programming models we consider. In Sect. 4 we summarize the classes of valid inequalities that are used as cutting planes in our implementations. Section 5 is dedicated to preprocessing routines that aim at tightening the given time windows, decompose the problem instance, or fix variables. In Sect. 6 we briefly describe our heuristics to obtain feasible solutions. Section 7 contains a description of other implementational details of the branch&cut algorithms. In Sect. 8 we report extensive computational results. Some concluding remarks are given in Sect. 9.
2. Related work

Time constrained sequencing and routing problems arise in many practical applications. Although the ATSP-TW is a basic model in many of these applications, not much attention has been paid to it so far. In most publications, exact algorithms play a minor role; the authors concentrate on the design of heuristics, often based on local search (see [39], among others).

For the symmetric TSP, to our knowledge, a few attempts have been made to solve problems with time windows to optimality, most of them based on implicit enumeration techniques (branch-and-bound, dynamic programming). Christofides et al. [15] describe a branch-and-bound algorithm in which the lower bound computation is performed via a state-space relaxation in a dynamic programming scheme. Solutions of problem instances of up to 50 nodes with “moderately tight” time windows are reported. Baker [8] also describes a branch-and-bound algorithm where the lower bound computations are reduced to the solution of a longest path problem on an acyclic network (the dual problem of a relaxation). The algorithm performs well on problems of up to 50 nodes when only a small percentage of the time windows overlap. Dumas et al. [20] present a dynamic programming algorithm for the TSP-TW that is able to solve problems of up to 200 nodes with “fairly wide” time windows. Here reductions of the state space and the state transitions are performed that are based on the time window structure. Balas and Simonetti [12] present a new dynamic programming algorithm that can be applied to a wide class of restricted travelling salesman problems. This approach yields good results on the ATSP-TW in case that the number of overlapping time windows is small [40]. Bianco et al. [14] present a dynamic programming algorithm for the TSP with time windows and precedence constraints and present computational results for instances up to 120 nodes. For surveys on time constrained routing and scheduling problems see [18, 19], among others.

Polyhedral approaches to solve problem instances to optimality are known to work well for the precedence constrained ATSP [7]. It is unclear whether a polyhedral approach can also handle time windows.

The ATSP-TW is related to the job-shop scheduling problem (JSSP) where one considers just one of the machines. Applegate and Cook [2] implemented a cutting plane algorithm for this problem type based on polyhedral investigations. Their computational results indicate that JSSP instances are difficult to solve – at least for this algorithmic approach.

Van Eijl [44] computationally compared two different formulations of the Delivery-Man Problem. As a by-product she obtained a branch-and-cut algorithm for the TSP-TW. Computational tests were performed on problem instances up to 15 nodes. She reported high running times.

It is not easy to compare the different approaches to the TSP-TW as there exist no standard benchmark problems. The instance sizes (typically expressed by the number of nodes) that can be solved depend extremely on the structure of the time windows. Moreover, in all the cases mentioned here, the authors only use randomly generated data. It is not clear what their findings mean for “real-world instances”. In general, authors reporting computational experiments with the TSP-TW conclude that the case where the time windows are active for about 50% is particularly difficult.
3. Notation and modelling

Given a node set $W \subseteq V$, let

$$A(W) := \{(i, j) \in A \mid i, j \in W\}$$

denote the set of all arcs with tail and head in $W$. For any two node sets $U, W \subseteq V$ let

$$(U : W) := \{(i, j) \in A \mid i \in U, j \in W\}$$

denote the set of arcs with tail in $U$ and head in $W$. To simplify notation, we write $(W : j)$ and $(j : W)$ instead of $(W : \{j\})$ and $(\{j\} : W)$, respectively. Given a node set $W \subseteq V$, we define

$$\delta^-(W) := \{(i, j) \in A \mid i \in V \setminus W, j \in W\},$$

$$\delta^+(W) := \{(i, j) \in A \mid i \in W, j \in V \setminus W\},$$

$$\delta(W) := \delta^-(W) \cup \delta^+(W).$$

The arc set $\delta(W)$ is called a cut. To simplify notation we write $\delta^-(v), \delta^+(v), \text{ and } \delta(v)$, instead of $\delta^-(\{v\}), \delta^+(\{v\}), \text{ and } \delta(\{v\})$, respectively.

For notational convenience, a path $P$ consisting of the arc set $\{(v_i, v_{i+1}) \mid i = 1, \ldots, k - 1\}$ is sometimes denoted by $P = (v_1, v_2, \ldots, v_k)$. If not stated differently, the path $P$ is always open and simple, i.e., $|P| = k - 1$ and $v_i \neq v_j$ for $i \neq j$. Moreover, we let

$$[P] := \{(v_i, v_j) \in A \mid 1 \leq i < j \leq k\}$$

denote the transitive closure of $P = (v_1, \ldots, v_k)$.

The minimal time delay for processing node $j$ immediately after node $i$ is given by

$$\vartheta_{ij} := p_i + c_{ij}.$$  

In the application that motivated this research the triangle inequality on $\vartheta$ is satisfied, i.e.,

$$\vartheta_{ij} \leq \vartheta_{ik} + \vartheta_{kj}, \text{ for all } i, j, k \in V. \quad (3.1)$$

If not stated differently we will assume throughout the paper that (3.1) holds.

Given a path $P = (v_1, \ldots, v_k)$, the earliest arrival time $t_{v_i}$ at node $v_i$ ($i = 1, \ldots, k$) along $P$ is computed as

$$t_{v_1} := r_{v_1}$$

$$t_{v_i} := \max\{t_{v_{i-1}} + \vartheta_{v_{i-1}v_i}, r_{v_i}\} \text{ for } i = 2, \ldots, k$$

This formula yields a waiting time $w_{v_i} := \max\{0, r_{v_i} - (t_{v_{i-1}} + \vartheta_{v_{i-1}v_i})\}$ which is positive whenever a node $v_i$ is reached before its release date. If $w_i = 0$ for all $i = 2, \ldots, k$, the path is called minimal. We denote by $\vartheta(P) := t_{v_k}$ the earliest arrival time at the last node of $P$. Notice that, for every minimal path $P$, one has $\vartheta(P) = r_{v_1} + \sum_{i=1}^{k-1} \vartheta_{v_{i-1}v_i}$.

To simplify notation, we sometimes write $\vartheta(v_1, v_2, \ldots, v_k)$ instead of $\vartheta(P)$ for $P = (v_1, v_2, \ldots, v_k)$.
A (Hamiltonian) tour \( T := (v_0, v_1, \ldots, v_n) \) of \( D \) starting at node \( v_0 = 0 \) at time \( r_0 = 0 \) is called \textit{feasible} if each node is visited within its time window, i.e., \( r_v \leq t_v \leq d_v \) for \( i = 1, \ldots, n \). A path \( P = (v_1, \ldots, v_k) \) with \( 2 \leq k \leq n \) is said to be \textit{infeasible} if it does not occur as a subpath in any feasible tour. Deciding whether a given path \( P \) is feasible is clearly an \( \mathcal{NP} \)-complete problem, even when \( P \) contains only one node, as in this case it amounts to deciding whether a feasible tour exists. Easily checkable and obvious sufficient conditions for infeasibility are given in the following lemma.

\( (3.2) \) Lemma. A given path \( P = (v_1, \ldots, v_k) \) is infeasible, if at least one of the following conditions holds:

(i) \( P \) violates the deadline for its last node \( v_k \), i.e., \( \vartheta(P) > d_v \).

(ii) The triangle inequality \((3.1)\) on \( \vartheta \) is satisfied and there exists a node \( w \) not covered by \( P \) such that both paths \( P_1 = (w, v_1, \ldots, v_k) \) and \( P_2 = (v_1, \ldots, v_k, w) \) violate the given deadline on their last node, i.e., \( \vartheta(P_1) > d_v \) and \( \vartheta(P_2) > d_w \).

In case condition (ii) above is satisfied, we say that node \( w \) cannot be \textit{covered by (an extension of) path \( P \).} If the triangle inequality on \( \vartheta \) is not satisfied condition (ii) can easily be modified by considering the \( \vartheta \)-shortest paths from \( w \) to \( v_1 \) and from \( v_k \) to \( w \) instead of \( P_1 \) and \( P_2 \).

Time windows induce precedences among the nodes. For example, whenever the \( \vartheta \)-shortest path from \( j \) to \( i \) is longer than \( d_i - r_j \) we can conclude that \( i \) has to precede \( j \) in any feasible solution. Then, let \( i \prec j \) denote the fact that \( i \) has to precede \( j \) in any ATSP-TW solution and let \( G_P := (V, R) \) denote the \textit{precedence digraph} where each arc \( (i, j) \in R \) represents a precedence relationship \( i \prec j \). Without loss of generality we may assume \( G_P \) to be acyclic and transitively closed. Moreover, let

\[
\begin{align*}
\pi(v) & := \{ i \in V | (i, v) \in R \}, \\
\sigma(v) & := \{ j \in V | (v, j) \in R \}.
\end{align*}
\]

represent the set of the \textit{predecessors} and \textit{successors} of a node \( v \in V \), respectively.

In the sequel we introduce three different integer programming models of the ATSP-TW, each defined on a different variable set. The first model involves binary arc variables \( x_{ij} \) as well as integer node variables \( t_i \). The second model uses only binary arc variables \( x_{ij} \), whereas the third one uses binary arc variables \( x_{ij} \) as well as integer arc variables \( y_{ij} \). For all the models, the binary variables \( x_{ij} \) for each arc \( (i, j) \in A \) can be interpreted as follows:

\[
x_{ij} := \begin{cases} 1, & (i, j) \in A \text{ is used in the tour,} \\ 0, & \text{otherwise.} \end{cases}
\]

A formal definition of the other variables will be given in the appropriate section.

3.1. Model 1

Miller, Tucker, and Zemlin [33] proposed to substitute the subtour elimination constraints for the TSP by a smaller class of inequalities and by introducing extra variables
These inequalities offer the advantage that they can easily be modified to take further side constraints into account (see [17]). For the ATSP with time windows, the $t$–variables are interpreted as time variables representing the arrival times at nodes, and the corresponding MTZ–inequalities can be written as

\[ t_i + \vartheta_{ij} - (1 - x_{ij}) \cdot M \leq t_j \quad i, j = 1, \ldots, n, i \neq j \]

(3.3)

where $M$ is a large real value. The ATSP-TW can therefore be formulated as an integer linear program as follows:

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad x(\delta^+(i)) = 1 \quad \forall i \in V \cup \{0\} \\
& \quad x(\delta^-)(i) = 1 \quad \forall i \in V \cup \{0\} \\
& \quad t_i + \vartheta_{ij} - (1 - x_{ij}) \cdot M \leq t_j \quad \forall (i, j) \in A, j \neq 0 \\
& \quad t_i \leq d_i \quad \forall i \in V \\
& \quad t_i \geq r_i \quad \forall i \in V \\
& \quad x_{ij} \in \{0, 1\} \quad \forall (i, j) \in A.
\end{align*}
\]

(3.4)

Where $N := \{0, 1, 2, \ldots\}$ denotes the set of nonnegative integers.

We will denote by

\[ p_1^{TW} := \text{conv}\{(x, t) \in \mathbb{R}^A \times V \mid (x, t) \text{ satisfies conditions (1)–(7) in (3.4)}\} \]

the ATSP–TW polytope based on Model 1. Note that instead of a global “big $M$”, an individual “big $M_{ij}$” may be defined for each inequality in (4), satisfying

\[ M_{ij} \geq d_i + p_i + c_{ij} - r_j. \]

It is easy to see that, for every feasible solution $(x, t)$ of (1)–(7), $x$ is the incidence vector of a feasible tour satisfying all given time windows.

This model has some disadvantages. First the MTZ–inequalities (3.3) are not very strong and they can be lifted in several ways (see Sect. 4). Furthermore, it is known from practical experience that a “big $M$”–modelling will cause computational problems. Our computations reported in Sect. 8 confirm this observation, even if some effort is spent on trying to reduce the big–M values involved. Even more important is that the $x$– and $t$–variables are only weakly linked via the MTZ–inequalities, i.e., the structure of the time windows has only very limited influence on the tour described by the $x$–variables.

### 3.2. Model 2

In a companion paper [4] we introduced a new model that is defined on binary arc variables only. In this model the time window restrictions are modelled by an additional
class of inequalities, the so–called infeasible path constraints. Let \( x(P) \) denote the sum of the variables corresponding to a path \( P \), i.e., \( x(P) := \sum_{i=1}^{k-1} x_{v_iv_{i+1}} \).

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad x(\delta^+(i)) = 1 \quad \forall i \in V \cup \{0\} \\
& \quad x(\delta^-(i)) = 1 \quad \forall i \in V \cup \{0\} \\
& \quad x(A(W)) \leq |W| - 1 \quad \forall W \subset V \cup \{0\}, 2 \leq |W| \leq n \\
& \quad x(P) \leq |P| - 1 = k - 2 \quad \forall \text{ infeasible path } P = (v_1, v_2, \ldots, v_k) \\
& \quad x_{ij} \in [0, 1] \quad \forall (i, j) \in A.
\end{align*}
\]

(3.5)

Inequalities (3.5)(4) forbid infeasible paths, i.e., paths violating the given time windows. Therefore, each solution \( x \) of (3.5)(1)–(5) is the incidence vector of a feasible Hamiltonian tour, and vice versa. The formulation of the infeasible path constraints as stated in (3.5)(4) can be very weak. In Sect. 4 we present several inequalities stronger than those of type (3.5)(4).

In analogy with Model 1, we denote by

\[ P_{TW}^2 := \text{conv}\{x \in \mathbb{R}^A \mid x \text{ satisfies conditions (1)–(5) in (3.5)}\} \]

the convex hull of all feasible solutions of Model (3.5).

### 3.3. Model 3

Maffioli and Sciomachen [32] and van Eijl [44] proposed a different model avoiding the need of “big \( M \)” terms. They introduced \(|A|\) additional integer arc variables \( y_{ij} \) with the property that \( x_{ij} = 0 \) implies \( y_{ij} = 0 \). If \( x_{ij} = 1 \) then \( y_{ij} \) denotes the time when the processing of node \( i \) is started and indicates that node \( j \) is processed after node \( i \). ATSP-TW can then be formulated as follows:

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad x(\delta^+(i)) = 1 \quad \forall i \in V \cup \{0\} \\
& \quad x(\delta^-(i)) = 1 \quad \forall i \in V \cup \{0\} \\
& \quad \sum_{i \neq j} y_{ij} + \sum_{i \neq j} \delta_{ij} \cdot x_{ij} \leq \sum_{k \neq j} y_{kj} \quad \forall j \in V \\
& \quad r_i \cdot x_{ij} \leq y_{ij} \leq d_i \cdot x_{ij} \quad i, j = 0, \ldots, n, i \neq j, i \neq 0 \\
& \quad x_{ij} \in [0, 1] \quad \forall (i, j) \in A.
\end{align*}
\]

(3.6)

As for the previous models we denote by

\[ P_{TW}^3 := \text{conv}\{(x, y) \in \mathbb{R}^A \times A \mid (x, y) \text{ satisfies conditions (1)–(5) in (3.6)}\} \]

the convex hull of all feasible solutions of Model (3.6).

Of course, the projection of \( P_{TW}^1 \) and \( P_{TW}^3 \) on the \( x \)-variables is the polytope \( P_{TW}^2 \). However, we do not have a description of any of these polytopes by linear equations and inequalities. Our goal is to find out how the LP relaxations of these polytopes that we know help to solve the ATSP-TW computationally.
4. Classes of valid inequalities

We summarize those known classes of inequalities for $P_{1}^{TW}$, $P_{2}^{TW}$, and for $P_{3}^{TW}$, that are used in at least one of our implementations.

Infeasible Path Elimination Constraints. Inequalities of this type forbid certain sub-paths that are infeasible, i.e., violate the given time windows. For a given infeasible path $P = (v_1, \ldots, v_k)$ the basic version of these inequalities is $x(P) \leq |P| - 1$. There exist, however, several possibilities to strengthen these inequalities, some of which were introduced in our companion paper [4]. Here are those we use in our actual implementation.

For every infeasible simple paths $P = (v_1, \ldots, v_k)$, the tournament constraint

\[ x([P]) := \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} x_{v_iv_j} \leq k - 2 \quad (= |P| - 1) \quad (4.1) \]

is valid for $P_{i}^{TW}$, $i = 1, 2, 3$. Obviously, if $A((v_1, \ldots, v_k))$ does not contain any feasible path the inequality can be strengthened to

\[ x(A((v_1, \ldots, v_k))) \leq k - 2. \quad (4.2) \]

Given a node set $W \subseteq V$, let $\Phi(W)$ denote a generic permutation of the nodes in $W$. For each node set $Q := \{v_1, \ldots, v_{k-1}\} \subset V$ and each node $v_k \in V \setminus Q$ such that all the paths of the form $(\Phi(Q), v_k)$ are infeasible, the inequality

\[ x(A(Q)) + x(Q : v_k) \leq k - 2 \quad (= |Q| - 1) \quad (4.3) \]

is valid for $P_{i}^{TW}$, $i = 1, 2, 3$. Note that a similar inequality can be defined for the case in which all the paths of the form $(v_1, \Phi(Q))$ are infeasible, namely

\[ x(v_1 : Q) + x(A(Q)) \leq k - 2 \quad (= |Q| - 1). \quad (4.4) \]

Moreover, for each node set $S := \{v_2, \ldots, v_{k-1}\} \subset V$ and for any two nodes $v_1, v_k \in V \setminus S$, $v_1 \neq v_k$, such that all paths of the form $(v_1, \Phi[S], v_k)$ are infeasible, the inequality

\[ x(v_1 : S) + x(A(S)) + x(S : v_k) + x_{v_1v_k} \leq k - 2 \quad (= |S|) \quad (4.5) \]

is valid for $P_{i}^{TW}$, $i = 1, 2, 3$. Note that inequalities (4.2), (4.3) and (4.4) are strengthenings of the subtour elimination constraint $x(A(Q)) \leq |Q| - 1$ as well as of the tournament constraints (4.1). However, it is not easy to decide whether all the paths of the form $(\Phi(Q), v_k)$, $(v_1, \Phi(Q))$, and $(v_1, \Phi[S], v_k)$ are infeasible, as required for the validity of inequalities (4.3)–(4.5). Easily checkable sufficient conditions are given by the next lemma.

(4.6) Lemma. Assume that the triangle condition (3.1) holds.

(a) Take any $Q \subset V$ and $v_k \in V \setminus Q$. If

\[ \min_{v_i \in Q} \{r_{v_i}\} + \sum_{v_i \in Q} \min_{v_j \in Q} \{|v_j \cup \{v_k\}| \} > d_{v_k} \]

then every path of the form $(\Phi(Q), v_k)$ is infeasible.
(b) Take any \( Q \subset V \) and \( v_1 \in V \setminus Q \). If
\[
r_{v_1} + \sum_{v_j \in Q} \min \{\vartheta_{v_j v_i} \mid v_i \in Q \cup \{v_1\}\} > \max_{v_i \in Q} \{d_{v_i}\}
\]
then every path of the form \((v_1, \Phi, Q)\) is infeasible.

(c) Take any \( S \subset V \) and \( v_1, v_k \in V \setminus S \), \( v_1 \neq v_k \). If
\[
r_{v_1} + \min \{\vartheta_{v_1 v_j} \mid v_j \in S\} + \sum_{v_i \in S} \min \{\vartheta_{v_i v_j} \mid v_j \in S \cup \{v_k\}\} > d_{v_k}
\]
then every path of the form \((v_1, \Phi, S, v_k)\) is infeasible.

A more involved generalization of tournament constraints can be obtained along the following lines. Suppose we are given a family \( \mathcal{P} := \{P_1, P_2, \ldots, P_k\} \) of node–disjoint simple paths, and let \( \omega \) be any permutation of the indices of \( \mathcal{P} \). The path \( P = (P_{\omega(1)}, P_{\omega(2)}, \ldots, P_{\omega(k)}) \) is called a concatenation of the paths in \( \mathcal{P} \). Now it may happen that the paths \( P_1, P_2, \ldots, P_k \) are feasible in themselves but that there is no way to concatenate them in a feasible way. In this case the inequality
\[
\sum_{i=1}^k x([P_i]) \leq \sum_{i=1}^k |P_i| - 1 \tag{4.7}
\]
is valid for \( P^{TW}_i \), \( i = 1, 2, 3 \).

**Lifted t-bounds.** It was observed in [17] that the bounds on the \( t \)-variables \( r_i \leq t_i \leq d_i \) can be strengthened by taking other arc combinations into account. Indeed, let \( a_{ji} := \max\{0, r_j - r_i + \vartheta_{ji}\} \) and \( b_{ji} := \max\{0, d_i - d_j + \vartheta_{ij}\} \). Then the inequalities
\[
(i) \quad r_i + \sum_{j=1}^n a_{ji} x_{ij} \leq t_i \quad \forall \ i \in V \\
(ii) \quad d_i - \sum_{j=1}^n b_{ij} x_{ji} \geq t_i \quad \forall \ i \in V \tag{4.8}
\]
are valid for \( P^{TW}_i \).

**Two–job cuts.** The ATSP-TW is related to the one–machine scheduling problem with time windows, i.e., the problem of sequencing \( n \) jobs on a single machine subject to a given set of time windows. Balas [9] and Dyer and Wolsey [21] considered the case where only release dates are present. The inequalities they derived can be used for the ATSP-TW too. At present, we only use the so-called two–job cuts (introduced by Balas [9]) in our implementation.

Suppose \( r_j < r_i + p_i \) and \( r_i < r_j + p_j \). The two–job cut involving order–dependent processing times can be written as
\[
(\vartheta_{ij} + r_i - r_j) \cdot t_i + (\vartheta_{ji} + r_j - r_i) \geq \vartheta_{ij} \cdot \vartheta_{ji} + r_i \cdot \vartheta_{ji} + r_j \cdot \vartheta_{ij} \tag{4.9}
\]
Violated inequalities of this class can be found by enumeration of all \( i \) and \( j \) satisfying \( r_j < r_i + p_i \) and \( r_i < r_j + p_j \).
TSP inequalities. Obviously, all valid inequalities for the (asymmetric) travelling salesman polytope \( P^T_n \) are also valid for the ATSP-TW. The classes of inequalities that we use in our implementation are:

- \( D^-_k \)-inequalities [25, 27, 22]
  \[
  \sum_{j=1}^{k-1} x_{ij_i j+1} + x_{ik_l} + 2 \sum_{j=3}^{k} x_{ij_i j} + \sum_{j=4}^{k-1} \sum_{h=3}^{j} x_{ij_i h} \leq k - 1 (4.10)
  \]

- \( D^+_k \)-inequalities [25, 27, 22]
  \[
  \sum_{j=1}^{k-1} x_{ij_i j+1} + x_{ik_l} + 2 \sum_{j=3}^{k-1} x_{ij_i j} + \sum_{j=4}^{k-1} \sum_{h=3}^{j-1} x_{ij_i h} \leq k - 1 (4.11)
  \]

- SD-inequalities [10]
  Given a handle \( H \subset V \), disjoint teeth \( T_1, \ldots, T_t \) odd, such that \( |T_i \cap H| = 1 \) and \( |T_i \setminus H| = 1 \), and (possibly empty) disjoint node sets \( S \) and \( D \), where \( (S \cup D) \subset V \setminus (H \cup T_1 \cup \ldots \cup T_t) \) and \( |S| + |D| = t \) is odd, the SD-inequalities have the form:
  \[
  x(\left((S \cup H) : (D \cup H)\right)) + \sum_{i=1}^{t} x(A(T_i)) \leq |H| + \frac{|S| + |D| + t - 1}{2}. (4.12)
  \]

- 2–matching constraints [27]
  Given vertex sets \( H, T_1, T_2, \ldots, T_k \subset V \), \( k \geq 3 \) and odd satisfying
  
  (i) \( |H \cap T_i| = 1 \) for \( i = 1, \ldots, k \),
  
  (ii) \( |T_i \setminus H| = 1 \) for \( i = 1, \ldots, k \),
  
  (iii) \( T_i \cap T_j = \emptyset \) for \( 1 \leq i < j \leq k \),

  the 2–matching constraint is given by
  \[
  x(A(H)) + \sum_{i=1}^{k} x(A(T_i)) \leq |H| + \frac{k - 1}{2}. (4.13)
  \]

SOP–inequalities. The precedence–constrained ATSP, also known as Sequential Ordering Problem (SOP), is a relaxation of the ATSP-TW. All valid inequalities for the SOP are therefore valid for the ATSP-TW. We summarize the classes of inequalities actually used in our implementation.

- Predecessor inequality (\( \pi \)-inequality) [11]
  Let \( S \subseteq V, \overline{S} := V \setminus S \). Then
  \[
  x((S \setminus \pi(S)) : (\overline{S} \setminus \pi(S))) \geq 1 \]  
  \[
  (4.14)
  \]
  is valid with respect \( P^T \_i, i = 1, 2, 3 \).
Solving the Asymmetric Travelling Salesman Problem with time windows by branch-and-cut

• Successor inequality ($\sigma$–inequality) [11]
  Let $S \subseteq V, S := V \setminus S$. Then
  \[ x((S \setminus \sigma(S)) : (S \setminus \sigma(S))) \geq 1 \] (4.15)
  is valid with respect $P^T_W, i = 1, 2, 3$.

• Predecessor–successor inequality ($\pi, \sigma$–inequality) [11]
  Let $X, Y \subseteq V$, s.t. $i \prec j \forall$ pairs $i \in X, j \in Y, W := \pi(X) \cup \sigma(Y)$. Then for all $S \subseteq V, s.t. X \subseteq S, Y \subseteq S$
  \[ x((S \setminus W) : (S \setminus W)) \geq 1 \] (4.16)
  is valid with respect $P^T_W, i = 1, 2, 3$.

• Precedence cycle breaking inequalities (pcb–inequality) [11]
  Let $S_1, \ldots, S_m \subseteq V, m \geq 2$, be disjoint node sets such that $\sigma(S_i) \cap S_{i+1} \neq \emptyset$ with $S_{m+1} := S_1$. Then
  \[ \sum_{i=1}^{m} x(A(S_i)) \leq \sum_{i=1}^{m} |S_i| - m - 1 \] (4.17)
  is valid with respect $P^T_W, i = 1, 2, 3$.

• Simple pcb–inequality
  The precedence cycle breaking inequality in its simplest form ($m = 2$ and $|S_2| = 1$)
  \[ x(A(S_1)) \leq |S_1| - 2. \] (4.18)

• “Special” inequalities [11]
  Let $S_1, S_2, S_3 \subseteq V \setminus \{1, n\}$ be disjoint node sets, with $S_1 \cap S_2 \neq \emptyset, S_2 \cap S_3 \neq \emptyset$.
  The following inequalities are valid with respect to $P^T_W, i = 1, 2, 3$:
  \[ \sum_{i=1}^{2} x(A(S_i)) + x(S_2 : S_1) \leq |S_2| + |S_1| - 2, \] (4.19)
  \[ \sum_{i=1}^{3} x(A(S_i)) + x(S_1 : S_3) \leq |S_1| + |S_2| + |S_3| - 3. \] (4.20)

**Strengthened ($\pi, \sigma$–inequalities).** In a companion paper [4] we introduced a strengthening of the ($\pi, \sigma$–inequalities originally introduced by Balas et al. [11] for the precedence–constrained ATSP.

Let $X$ and $Y$ be two disjoint node sets such that $i \prec j$ for all $i \in X$ and $j \in Y$, and define $W := \pi(X) \cup \sigma(Y)$. Assume that the triangle inequality (3.1) on $\vartheta$ is satisfied and define
\[ \bar{W} := W \cup \{k \in V \setminus (X \cup Y) \mid \exists i \in X \text{ and } j \in Y \text{ s.t. } \vartheta(i, k, j) > d_j\} \]
and
\[ Q := \{(u, v) \in \delta^+(S) \mid \exists i \in X \text{ and } j \in Y \text{ s.t. } \vartheta(i, u, v, j) > d_j\}. \]
Then for all $S \subseteq V$ such that $X \subseteq S$ and $Y \subseteq \overline{S}$ the inequality
\[ x((S \setminus \tilde{W}) : (\overline{S} \setminus \tilde{W})) \setminus Q \geq 1 \] (4.21)
is valid for $P_{TW1}$, $P_{TW2}$, and $P_{TW3}$.

The $\pi$– and $\sigma$–inequalities introduced by Balas et al. [11] can be strengthened as well (see Ascheuer [3]). Our computational experience showed, however, that for the problem instances in our test bed no such strengthenings were possible. Therefore, we omit these inequalities here.

**Strengthened MTZ-inequalities.** Desrochers and Laporte [17] observed that the MTZ–subtour elimination constraints (3.3) can be lifted by taking the reverse arcs $(j, i) \in A$ and infeasible arc combinations into account. Let $a_{ji} := \max\{\vartheta_{ji}, r_i - d_j\}$, $M_{ij} \geq d_i + \vartheta_{ij} - r_j$. Then for all $i, j = 1, \ldots, n, i \neq j$ the inequality
\[ t_i + \vartheta_{ij} - (1 - x_{ij}) \cdot M_{ij} + (M_{ij} - \vartheta_{ij} - a_{ji}) \cdot x_{ji} \leq t_j \] (4.22)
is valid for $P_{TW1}$.

In addition, suppose $i, j, k \in V$ are such that $r_k + \vartheta_{ki} + \vartheta_{ij} > d_j$, and choose values $M$ and $b_{ki}$, such that $M \geq \max_{ij} \{c_{ij} + c_{ji}\}$ and $b_{ki} \leq M - c_{ij} - \min\{d_k + c_{ki}, d_i\}$. Then the inequality
\[ t_i + \vartheta_{ij} - (1 - x_{ij}) \cdot M + (M - \vartheta_{ij} - a_{ji}) \cdot x_{ji} + b_{ki} x_{ki} \leq t_j \] (4.23)
is valid for $P_{TW1}$.

In case that precedences are present, the MTZ–inequalities can be further strengthened [3]. Assume $i \prec j$. As $i$ must be scheduled before $j$, we know that $t_i \leq t_j$ and, even more, that the inequality
\[ t_i + \vartheta_{ij} x_{ij} \leq t_j \] (4.24)
is valid. Note, that inequality (4.24) can be strengthened to
\[ t_i + \vartheta_{ij} x_{ij} \leq r_j \] (4.25)
in case that $d_i + \vartheta_{ij} \leq r_j$ holds.

To close this section we remark that the lifted $t$-bounds (4.8) and the strengthened MTZ-inequalities (4.22)–(4.25) apply only to Model 1, while all other inequalities mentioned above are valid for all three models.

5. **Data preprocessing**

As for many other combinatorial optimization problems, preprocessing is an important part of an efficient implementation. Its main aim is to construct a “tighter” equivalent formulation of the problem, such that no optimal solution of the original problem is lost and each optimal solution of the tighter problem corresponds to an optimal solution of the original problem.
For the ATSP-TW, preprocessing includes three main steps: tightening the time windows, constructing precedences among the nodes, and fixing variables permanently. In addition, we detect paths $P = (v_1, v_2)$ of length one which are infeasible due to the criteria given by Lemma (3.2). If such paths are detected, the corresponding arcs $(v_1, v_2)$ cannot be used in any feasible solution and are therefore deleted from the feasible arc set.

### 5.1. Tightening of the time windows

In this section we list some criteria (see, e.g., [16, 19]) that allow us to increase the release date (resp. to decrease the due date) of certain nodes.

#### 5.1.1. Release date adjustment

If the earliest arrival time at node $k \in V$ from any of its possible predecessors is bigger than its release date $r_k$ (see Fig. 5.1(a)), the release date of $k$ can be increased, i.e.,

$$ r_k := \max\{r_k, \min_{(i,k) \in A} \{r_i + \vartheta_{ik}\}\} \quad \forall k \in V \text{ s.t. } \delta^-(k) \neq \emptyset. \quad (5.1) $$

In order to avoid waiting times at the possible successor nodes of $k \in V$, the earliest possible starting time of $k$, and therefore its release date may be shifted (see Fig. 5.1(b)), i.e.,

$$ r_k := \max\{r_k, \min_{(k,j) \in A} \{r_j - \vartheta_{kj}\}\} \quad \forall k \in V \text{ s.t. } \delta^+(k) \neq \emptyset. \quad (5.2) $$

#### 5.1.2. Due date adjustment

If the due date $d_k$ of node $k \in V$ is larger than the latest possible arrival time at node $k$ from any of its predecessors, the due date may be decreased (see Fig. 5.1(c)), i.e.,

$$ d_k := \min\{d_k, \max_{(i,k) \in A} \{d_i + \vartheta_{ik}\}\} \quad \forall k \in V \text{ s.t. } \delta^+(k) \neq \emptyset. \quad (5.3) $$

If the latest possible departure time from node $k \in V$ in order to fulfill all time window constraints for its successors is smaller than its due date, then it can be decreased (see Fig. 5.1(d)), i.e.,

$$ d_k := \min\{d_k, \max_{(k,j) \in A} \{d_j - \vartheta_{kj}\}\} \quad \forall k \in V \text{ s.t. } \delta^+(k) \neq \emptyset. \quad (5.4) $$
5.2. Construction of precedences

Whenever the time windows for two nodes \( i \) and \( j \) are “non-overlapping” we can infer a precedence. E.g., if \( r_j + \vartheta_{ji} > d_i \) we know that node \( i \) has to precede node \( j \) in any feasible solution to the ATSP-TW. From time windows, therefore, precedences among the nodes in \( V \) can be derived and the methodology developed for the SOP [7] can also be applied to the ATSP-TW.

Let \( i \prec j \) denote the fact that node \( i \) has to precede node \( j \) in every feasible solution, and let \( P = (V, R) \) be the precedence digraph defined on the same node set as \( D \) and where an arc \( (i, j) \in R \) represents a precedence relationship \( i \prec j \). Clearly, \( P \) must be acyclic and can be assumed to be transitively closed.

5.3. Elimination of arcs

By construction, if \( (i, j) \) is in the arc set \( R \) of the precedence digraph, the arc \( (j, i) \) cannot be contained in any feasible Hamiltonian path. So we can delete all arcs \( (j, i) \) from \( A \) for which \( (i, j) \in R \). Furthermore, for any nodes \( i, j, k \in V \) with \( (i, j) \in R \) and \( (j, k) \in R \), we can conclude that arc \( (i, k) \) cannot be used in any feasible Hamiltonian path as node \( j \) has to be sequenced between \( i \) and \( k \). Therefore, those arcs \( (i, k) \) arcs are eliminated as well.

For all other arcs \( (i, j) \in A \), we start with the feasible path \( P = (i, j) \) and try to concatenate \( P \) with paths formed by nodes in a given node set \( Q := \{v_1, \ldots, v_k\} \). If all these concatenations result in an infeasible path, we know that arc \( (i, j) \) cannot be used in any feasible solution and it is eliminated from the feasible arc set \( A \). For sets \( Q \) of small cardinality this can be checked by enumerating all possible paths containing arc \( (i, j) \). In our implementation we only consider the case \(|Q| \leq 2\).

Table 5.1 shows the effect these three preprocessing steps (applied iteratively several times) have on some of our instances. E.g., about 50% of the arcs can be fixed or eliminated and lots of precedences are generated resulting in a considerable decrease of the model size.

6. Heuristics

Recall that it is an \( \mathcal{NP} \)-complete problem to find a feasible solution to the ATSP-TW. Therefore, unless \( \mathcal{P} = \mathcal{NP} \), there is no efficient procedure that is guaranteed to terminate with such a solution. As a consequence, we run several construction heuristics varying from very simple sorting heuristics to more sophisticated insertion heuristics. Whenever a feasible solution is found we run improvement heuristics in order to obtain a better tour.

For implementational convenience we split the depot node 0 and create an additional dummy node \( n + 1 \) such that \( i < n + 1 \ \forall i \in V \). Every feasible tour corresponds to a feasible Hamiltonian path starting at 0 and ending at \( n + 1 \) in the new digraph.
Solving the Asymmetric Travelling Salesman Problem with time windows by branch-and-cut

Table 5.1 Preprocessing results

|     | Iter. | R1 | R2 | TW1 | TW2 | TW3 | TW4 | Fix1 | Fix2 | |A|
|-----|------|----|----|-----|-----|-----|-----|------|------|-----|
| rbg10a | 110 | 2 | 27 | 29 | 0 | 3 | 0 | 0 | 0 | 0 | 54 |
| rbg10b | 110 | 10 | 32 | 19 | 25 | 0 | 5 | 0 | 30 | 0 | 41 |
| rbg16a | 272 | 3 | 3 | 94 | 98 | 0 | 3 | 4 | 0 | 0 | 79 |
| rbg16b | 272 | 3 | 54 | 49 | 0 | 1 | 1 | 0 | 0 | 2 | 167 |
| rbg17a | 306 | 39 | 62 | 26 | 235 | 0 | 4 | 0 | 74 | 5 | 174 |
| rbg19a | 380 | 3 | 154 | 153 | 0 | 1 | 3 | 0 | 0 | 2 | 71 |
| rbg19b | 380 | 2 | 90 | 79 | 0 | 2 | 1 | 0 | 0 | 0 | 211 |
| rbg19c | 380 | 39 | 74 | 36 | 245 | 0 | 4 | 0 | 72 | 0 | 229 |
| rbg20a | 420 | 3 | 94 | 98 | 0 | 3 | 4 | 0 | 1 | 79 |
| rbg27a | 756 | 46 | 142 | 438 | 0 | 3 | 0 | 92 | 15 | 487 |
| rbg48a | 2352 | 105 | 577 | 360 | 1639 | 0 | 6 | 0 | 247 | 0 | 1288 |
| rbg49a | 2450 | 114 | 734 | 580 | 1842 | 1 | 5 | 0 | 102 | 1 | 1083 |
| rbg50a | 2500 | 3 | 485 | 429 | 0 | 3 | 2 | 0 | 0 | 7 | 1629 |
| rbg50b | 2550 | 115 | 732 | 591 | 1964 | 1 | 4 | 0 | 102 | 0 | 1175 |
| rbg50c | 2550 | 108 | 611 | 450 | 1604 | 0 | 4 | 0 | 172 | 0 | 1396 |

| m     | Iter. | R1 | R2 | TW1 | TW2 | TW3 | TW4 | Fix1 | Fix2 | |A|
|-------|------|----|----|-----|-----|-----|-----|------|------|-----|
| rbg10a | 110 | 2 | 27 | 29 | 0 | 3 | 0 | 0 | 0 | 54 |
| rbg10b | 110 | 10 | 32 | 19 | 25 | 0 | 5 | 0 | 30 | 0 | 41 |
| rbg16a | 272 | 3 | 3 | 94 | 98 | 0 | 3 | 4 | 0 | 0 | 79 |
| rbg16b | 272 | 3 | 54 | 49 | 0 | 1 | 1 | 0 | 0 | 2 | 167 |
| rbg17a | 306 | 39 | 62 | 26 | 235 | 0 | 4 | 0 | 74 | 5 | 174 |
| rbg19a | 380 | 3 | 154 | 153 | 0 | 1 | 3 | 0 | 0 | 2 | 71 |
| rbg19b | 380 | 2 | 90 | 79 | 0 | 2 | 1 | 0 | 0 | 0 | 211 |
| rbg19c | 380 | 39 | 74 | 36 | 245 | 0 | 4 | 0 | 72 | 0 | 229 |
| rbg20a | 420 | 3 | 94 | 98 | 0 | 3 | 4 | 0 | 1 | 79 |
| rbg27a | 756 | 46 | 142 | 438 | 0 | 3 | 0 | 92 | 15 | 487 |
| rbg48a | 2352 | 105 | 577 | 360 | 1639 | 0 | 6 | 0 | 247 | 0 | 1288 |
| rbg49a | 2450 | 114 | 734 | 580 | 1842 | 1 | 5 | 0 | 102 | 1 | 1083 |
| rbg50a | 2500 | 3 | 485 | 429 | 0 | 3 | 2 | 0 | 0 | 7 | 1629 |
| rbg50b | 2550 | 115 | 732 | 591 | 1964 | 1 | 4 | 0 | 102 | 0 | 1175 |
| rbg50c | 2550 | 108 | 611 | 450 | 1604 | 0 | 4 | 0 | 172 | 0 | 1396 |

m : Number of arcs in the original input digraph
Iter. : Number of preprocessing loops
R1 : Number of precedence relationships among the nodes
R2 : Number of eliminated transitive relationships arcs
TW1 : Number of release date adjustments due to (5.1).
TW2 : Number of release date adjustments due to (5.2).
TW3 : Number of deadline adjustments due to (5.3).
TW4 : Number of deadline adjustments due to (5.4).
Fix1 : Number of variables fixed to 0 due to the criterion described in Sect. 5.3 (|Q| = 1).
Fix2 : Number of variables fixed to 0 due to the criterion described in Sect. 5.3 (|Q| = 2).
|A| : Number of remaining arcs / variables

6.1. Construction heuristics

**Sorting Heuristics:**
We apply the following sorting criteria:

1. Check if the trivial sequence (0, 1, 2, 3, ..., n − 1, n, n + 1) is feasible.
2. Sort the nodes according to increasing release dates and check whether this sequence is feasible.
3. Sort the nodes according to increasing due dates and check whether this sequence is feasible.
4. Sort the nodes according to increasing midpoints of the time windows \( m_i := r_i + \frac{d_i}{2} \) and check whether this sequence is feasible.

**Nearest-Feasible-Neighbor Heuristic:**
Starting with each feasible arc (0, i) ∈ A we run a nearest feasible neighbor heuristic, i.e., we enlarge the current subpath (0, v_1, v_2, ..., v_k) by an arc (v_k, v_l) resulting in the smallest increase in the objective value and guaranteeing feasibility.

**Insertion Heuristics:**
Starting with a shortest path in A from 0 to n + 1 we enlarge the current partial path \( P' := (0, v_1, ..., v_k, n + 1) \) by a node j satisfying a certain insertion criterion. Let \( W := V \setminus \{v_1, ..., v_k\} \) and \( d_{\min}(j) := \min[c_{v_j} + c_{jv_{j+1}} - c_{v_{j+1}}] i \in V \setminus W, \)
$v_l \in P'$ and $(0, v_1, \ldots, v_l, j, v_{l+1}, \ldots, v_k, n+1)$ is feasible. We apply the following strategies, each resulting in a different heuristic:

1. Among all unsequenced nodes choose the node $j \in W$ that causes the lowest increase in the path length, i.e., $d_{\text{min}}(j) = \min(d_{\text{min}}(l)| l \in W)$.

2. Among all unsequenced nodes choose the node $j \in W$ that has the lowest number of feasible insertion positions and insert this node at the cheapest of these positions.

The best solution found by any of the construction heuristics is passed to the improvement heuristics.

The time complexity of the heuristics mentioned above is dominated by sorting. For the problem sizes considered in this paper, the running times are almost immeasurable.

6.2. Improvement heuristics

**Swap Heuristic:**
Given a feasible tour $T = (v_0, v_1, \ldots, v_{n+1})$ we scan through the sequence and check whether swapping two subsequent nodes $v_i$ and $v_{i+1}, i = 1, \ldots, n-1$, results in a feasible solution with better objective value, in which case the swap is accepted. This procedure is repeated until no further improvement is achieved.

**Two-Node-Exchange Heuristic:**
This is a generalization of the Swap-heuristic where any two nodes (not only subsequent nodes) in the current sequence are exchanged. If this results in a better feasible tour the exchange is accepted. This procedure is repeated until no further improvement is achieved.

**Node-Reinsertion Heuristic:**
Given a feasible tour $(v_0, v_1, \ldots, v_i, v_j, v_k, \ldots, v_{n+1})$ and any inner node $v_j$, we construct a partial tour $T' = (v_0, v_1, \ldots, v_i, v_k, \ldots, v_{n+1})$ by deleting $v_j$. We then try to reinsert $v_j$ in the best position in $T'$ such that the new sequence $T$ is feasible. If this results in a better objective value we accept $T$ to be the new tour and repeat until no further improvement is achieved.

**Arc-Reinsertion Heuristic:**
Given a feasible tour $(v_0, v_1, \ldots, v_i, v_k, v_l, \ldots, v_{n+1})$ we construct a partial tour $T' = (v_0, v_1, \ldots, v_i, v_l, \ldots, v_{n+1})$ by deleting any two consecutive nodes $v_j$ and $v_k$. We then try to reinsert the pair $(v_j, v_k)$ at any position in $T'$ such that the new sequence $T$ is feasible. If this results in a better objective value we accept $T$ to be the new tour and repeat until no further improvement is achieved.

**Arc-Reversal Heuristic:**
Given a feasible tour $(v_0, \ldots, v_j, v_k, \ldots, v_m, \ldots, v_{n+1})$ we construct a tour $T := (v_0, \ldots, v_j, v_l, \ldots, v_k, \ldots, v_{n+1})$ by reversing the subpath $(v_k, \ldots, v_l)$ such that the new sequence $T$ is feasible. If this results in a better objective value we accept $T$ to be the new tour and repeat until no further improvement is achieved.

**Or-Exchange Heuristic:** (see as well [34,38,41])
Given a feasible tour $(v_0, v_1, \ldots, v_i, v_j, \ldots, v_{n+1})$ we remove the subpath $(v_i, \ldots, v_j)$ and try to reinsert it between any two subsequent nodes $v_l$ and $v_{l+1}$ such
that the new sequence $T$ is feasible. If this results in a better objective value we accept $T$ to be the new tour and repeat until no further improvement is achieved. We restrict ourselves to paths involving up to 5 nodes.

To our knowledge, there does not exist a proof of polynomiality for the swap, renewal, exchange or reinsertion heuristics we have considered. The running times for our problems in practice are marginal. None of these heuristics dominated the others and none was dominated by the remaining ones. That is why we used them all.

The heuristics are called in the following order:

(6.1) Initial Heuristics.

1. Run Sorting Heuristics
2. Run Nearest Feasible Neighbor Heuristics
3. Run Insertion Heuristic 1
4. Run Insertion Heuristic 2
5. IF no feasible sequence found STOP
6. DO until no further improvement is achieved
   (a) Run Or-Exchange-Heuristic
   (b) Run Arc-Reversal Heuristic
   (c) Run Swap Heuristic
   (d) Run Arc-Reinsertion Heuristic
   (e) Run Node-Reinsertion Heuristic
   (f) Run Two-Node-Exchange Heuristic

The best solution found by any of the procedures is the initial feasible solution passed to the branch–and–cut code. Our computational experiments showed that these heuristics are outperformed by the LP-based heuristic procedures described in the next section.

6.3. LP–exploitation heuristic

In order to make use of the information obtained during the branch–and–cut execution, an LP–exploitation heuristic is run after each LP–solution. Given the current fractional point $x^*$, we construct the digraph $D^* := (V, A^*)$ where $(i, j) \in A^*$ if and only if $x^*_{ij} > 0$. We execute two steps.

First, in case that $D^*$ is very sparse (in our current implementation this means that $|A^*| \leq 1.25 \cdot |V|$), we apply a branch–and–bound like implicit enumeration scheme, where we backtrack as soon as the current path becomes infeasible or the cost of the path is higher than the cost of the best feasible solution so far. We expected and observed that this procedure rarely finds a feasible tour when applied to the optimal solutions of the first LP relaxations. Feasible tours show up only at the end of the cutting plane procedure after sufficiently many inequalities have been added to the initial LP forcing out subtours, infeasible paths, and the like.

In a second step, that is always performed, we set up a modified problem instance by changing the costs $c_{ij}$ of all $(i, j) \in A$ as follows, $c^*_{ij} := (1 - x^*_i) \cdot c_{ij}$. With this modified
cost-matrix $C^*$ we run all of the following construction heuristics: Nearest–Feasible–
Neighbour, Insertion 1, and Insertion 2 heuristics. If a feasible solution is found we try to
improve it by applying one of the following improvement heuristics: Node Reinsertion,
Arc Reinsertion, Swap, and Arc Reversal. The improvement heuristics are called with
the original costs $c_{ij}$. In order to avoid calling the improvement heuristics more than
once with the same input solution we use an hash table in which the insertion key is
the value of the input solution along with the index of the improvement heuristic. If the
heuristic was already called with a solution of the same value, it is skipped. Note that
different solutions may have the same value. Hence, we may miss a solution that leads
to an improvement. Nevertheless, this strategy led to a dramatic reduction in computing
time needed for the LP–exploitation heuristics.

7. The branch&cut algorithm

We implemented four different branch&cut–algorithms: One for each Model (3.4)–(3.6),
plus an advanced implementation based on Model 2 using the branch&cut framework
ABACUS [30,42]. We performed the advanced implementation only for this model as
the preliminary version of ABACUS to which we had access supported binary variables
only.

We assume from now on that the preprocessing steps described in Sect. 5 have been
performed resulting in a reduced digraph $D = (V, A)$ and tightened time windows.

7.1. Initial linear program

The initial linear programs for the three models are generated as follows.

Model 2. This model involves only the variables $x_{ij}$, $(i, j) \in A$. We generate the vari-
bables $x_{ij}$ that correspond to arcs $(i, j)$ in the 5–nearest–neighbors digraph and to the
arcs of the best feasible tour found by the initial heuristics. (The remaining variables are
later taken into account by pricing). We generate the nonnegativity constraints and the
degree constraints

$$x(\delta^-(i)) = 1 \quad \forall i \in V \cup \{0\}$$
$$x(\delta^+(i)) = 1 \quad \forall i \in V \cup \{0\}.$$ 

restricted to the initial set of $x_{ij}$-variables.

Model 1. We generate the same $x_{ij}$-variables and constraints as in Model 2. We add
all node variables $t_i$, $i \in V$, and the corresponding lifted $t$–bounds (4.8). We generate,
for every variable $x_{ij}$, the MTZ–inequality (3.4)(3), and, whenever possible, we add
strengthened MTZ–inequalities of type (4.22), (4.23), (4.24).

Model 3. We generate the same $x_{ij}$-variables and constraints as in Model 2 and, for
each $x_{ij}$-variable, the corresponding variable $y_{ij}$. We add the inequalities (3.6)(3) and
(3.6)(4).
7.2. Separation routines

In this section we describe the separation procedures for the classes of inequalities listed in Sect. 4. For some of these classes we use routines described in the literature.

- **Subtour Elimination constraints (SEC):** We use the separation routine developed by Padberg and Rinaldi [35]. Whenever a violated SEC is found, we check if it can be strengthened to a \( \pi-, \sigma-, \) or PCB-inequality ((4.14), (4.15), or (4.17)).

- **2-matching constraints:** We use a heuristic separation procedure as described in [26].

- **\( \pi-, \sigma-, (\pi, \sigma)-\)inequalities:** We use the heuristic separation procedure for the “weak” version of these inequalities, as proposed in [11].

- **“Special” inequalities (4.19) and (4.20), PCB–inequalities (4.17):** We use the “shrinking procedure” proposed in [11]. We shrink saturated SECs both in the LP-solution digraph \( D^* \) and the precedence digraph. If an infeasible arc (i.e., a reverse or transitive arc of a precedence arc) or a cycle in the precedence digraph is detected we have a violated pcb–inequality or a violated inequality of type (4.19) or (4.20).

- **\( D_k-\)inequalities:** We use the separation procedure described in [23].

- **SD–inequalities:** We use the separation procedure described in [23].

- **Infeasible Path Elimination Constraints (IPEC):** In the test phase of our implementation we often observed the following: As soon as an IPEC is generated, the LP–solution tries to react to this cutting plane by taking a short “detour” or a “short cut”. Therefore, in the final implementation we first check whether a minor modification of an already–generated infeasible path constraint is violated. If these trivial separation checks are not successful, tournament constraints (4.1) are separated with the help of the following simple enumeration procedure. Suppose we are given a (fractional) point \( x^* \). It can be shown (Savelsbergh [37]) that there are only polynomially many paths \( P_k \) for which \( \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} x^*_w v - k + 2 \) is greater than 0, under the assumption that \( \sum_{i \in V} x^*_w v = \sum_{j \in V} x^*_v w = 1 \). These paths can easily be detected by enumeration (backtrack as soon as \( \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} x^*_w v - k + 2 \leq 0 \)). If a violated tournament constraint is found, we check, whether it can be lifted to an inequality of type (4.2)–(4.5). If no cut is found, we verify if a concatenation of paths corresponding to variables \( x^*_w v = 1 \) creates an infeasibility, i.e., if an inequality of type (4.7) is violated.

- **Strengthened (\( \pi, \sigma \))–inequalities:**

  If \( X = \{i\} \) and \( Y = \{j\} \), the strengthened \( (\pi, \sigma) \)–inequality is called simple. The separation problem for the simple strengthened \((\pi, \sigma) \)–inequalities can heuristically be solved via a separation procedure similar to the one used for the weak \((\pi, \sigma) \)–inequalities (Balas et al. [11]). Given the LP fractional point \( x^* \), we set up a capacitated digraph \( D^* := (V, A^*) \), \( A^* := \{(i, j) \in A \mid x^*_w v > 0\} \). To each \( (i, j) \in A^* \) associate a capacity \( c^*_w v := x^*_w v \). For all \( i < j \) we then apply the following procedure. We construct a digraph \( \tilde{D} := (\tilde{V}, \tilde{A}) \) from \( D^* \) by deleting
  - all nodes in \( \pi(i) \cup \sigma(j) \),
  - all nodes \( k \) such that \( (i, k, j) \) is infeasible, i.e., \( r_i + \theta_{ik} + \theta_{kj} > d_j \),
  - all arcs \((u, v)\) such that \((i, u, v, j)\) is infeasible, i.e., \( r_i + \theta_{iu} + \theta_{uv} + \theta_{vj} > d_j \).
If we do not succeed in sending one unit of flow from $i$ to $j$ in $\tilde{D}$, then the minimum capacity cut in $\tilde{D}$ separating $i$ from $j$ defines a violated simple generalized $(\pi, \sigma)$--inequality.

- **Two-Job-Cuts:** Violated inequalities of this class can be found by enumeration of all $i$ and $j$ satisfying $r_j < r_i + p_i$ and $r_i < r_j + p_j$.

- **Pool-Separation:** During the run of the algorithm we maintain a pool of active and nonactive valid inequalities. An inequality is called active if it is both stored in the constraint matrix of the current LP and in the pool, whereas inequalities that are only stored in the pool are called nonactive. The pool is initially empty. Each generated cut is added to the constraint matrix and is stored in the pool. As soon as a constraint is nonbinding in the current LP it becomes inactive, i.e., it is removed from the constraint matrix but is still kept in the pool. The inequalities in the pool can be used either to reconstruct linear programs from scratch, or to check if any of the cuts generated in an earlier iteration of the algorithm is violated by the actual LP--solution (pool separation).

**Separation-Order.** The separation routines are called in the following order:

1. Pool Separation.
2. Subtour elimination constraints.
3. “Shrinking–procedure”.
4. $\pi$--inequalities.
5. $\sigma$--inequalities.
6. Infeasible Path Elimination Constraints.
7. $D_k$--inequalities.
8. SD--inequalities.
9. Strengthened $(\pi, \sigma)$--inequalities.
10. 2--matching constraints.
11. Two–Job Cuts (only for Model 1).

Whenever one of the procedures generates a cutting plane all subsequent routines are skipped.

**7.3. Variable fixing**

Within the branch&cut tree (BC--tree, for short) we apply two types of variable fixings. We say that a variable $x_{ij}$ is fixed to its upper or lower bound, if this operation is valid for the whole BC--tree. If this is valid only for the current BC--node and all of its children-nodes we say that this node is set to its upper or lower bound.

If a variable is set (resp. fixed) to zero, the corresponding arc is deleted from the current feasible arc set $A$. Note that fixing is a permanent operation, but due to the local character of variable settings the feasible arcs may differ between different nodes of the BC-tree.

**7.3.1. Logical implications.** Due to the fixing (resp. setting) of a variable, further fixings (resp. settings) may be performed. First, assume that variable $x_{ij}$ was set (resp. fixed) to zero. If after this operation either $|\delta^+(i)| = 1$ or $|\delta^-(j)| = 1$, we know that the arc
leaving $i$ or entering $j$ has to be used in any feasible tour. Therefore, the corresponding variable can be set (resp. fixed) to one.

Now assume, that variable $x_{ij}$ was fixed to one. An update of the time windows is possible, i.e., we set

$$r_j := \max\{r_j, r_i + \vartheta_{ij}\}$$
$$d_i := \min\{d_i, d_j - \vartheta_{ij}\}$$

and call the preprocessing routines described in Sect. 5. In order to guarantee that all infeasible path constraints, that are based on the time windows, remain globally valid throughout the BC-tree, this operation is only applied when a variable is fixed (not when it is set) to 1.

In addition, the following reductions can be applied:

$$x_{ij} = 1 \Rightarrow x_{ji} = 0$$
$$x_{ik} = 0 \quad \forall k \in V \setminus \{j\}$$
$$x_{kj} = 0 \quad \forall k \in V \setminus \{i\}$$

$$x(i : \pi(j)) = 0$$
$$x(\sigma(j) : i) = 0$$
$$x(\pi(i) : \sigma(j)) = 0$$
$$x(\pi(j) : \sigma(i)) = 0.$$

### 7.3.2. Reduced cost fixing

Nonbasic active variables can be fixed to their current value using reduced cost criteria. To this end, assume that we are given a global upper bound $\text{gub}$ and a global lower bound $\text{glb}$. For a nonbasic variable $x_{ij}$, we compute the associated reduced cost $\text{rij}$. In case that $x_{ij} = 0$ and $\text{glb} + \text{rij} > \text{gub}$ we are allowed to fix variable $x_{ij}$ to zero, whereas $x_{ij} = 1$ and $\text{glb} - \text{rij} > \text{gub}$ implies that $x_{ij}$ can be fixed to one. Using the same arguments nonbasic active variables can be set to their current value using a local lower bound and reduced costs instead of the global ones.

### 7.4. Further implementation details

**Branching.** Branching is performed on $x-$variables only. The branching variable $x_{ij}$ is chosen to be one that is closest to 0.5. If there exist several such variables, one with highest cost coefficient $c_{ij}$ is chosen.

**Enumeration strategy.** For the comparison of the three models, Depth–First–Search is applied as enumeration strategy. ABACUS also supports Breadth–First–Search and Best–First–Search. Computational test have shown that Best–First–Search yields slightly better results than the other strategies. Therefore, we have chosen Best–First–Search as default strategy for the advanced implementation of Model 2.

**Pricing frequency:** Nonactive variables are priced out at each 5-th LP-solution.

**Tailing off:** In a certain BC–node, whenever new cuts are added but the increase in the objective function is not sufficiently large, we say that “tailing off” occurs and perform a branching step. In our implementation we resort to branching whenever the last 10 LPs produced no improvement in the lower bound, or in case an improvement of only 1% has been achieved in the last 20 LPs.
8. Computational results

All implementations were done in the programming language C on a SUN SPARC Station 10 and compiled with the Workshop Compiler cc 4.2 with options -fast -Xc. Only for the advanced implementation based on Model 2 we used a preliminary version of the general purpose branch&cut-framework ABACUS [29–31], supporting 0/1-variables only. Therefore, the comparison of the three different models was performed on ad hoc implementations we developed.

The LPs were solved using the callable library of CPLEX 4.09. The codes were tested on a set of real-life data from a joint project with industry that had the aim to minimize the unloaded travel time of a stacker crane within an automatic storage system (see Ascheuer [3] for details). This set of test problems was enlarged by instances for which we relaxed some randomly selected time windows. These datasets are based on the 27–node real-life instance rgb027a and are denoted by rgb27.b.x, where x gives the number of nodes for which time windows are present.

8.1. Comparison of the three models

In a first phase we performed computational experiments with all three models of the ATSP-TW described in Sect. 3. Our aim was to gather from these test runs information that would allow us to choose a “winner”, i.e., a model that, for the range of problem instances we address, displays the best computational performance in practice.

From the test runs we have made, we have chosen to present 11 benchmark problems that cover our range of typical instance sizes and that all three models were able to solve to optimality. The data of these test runs are, by and large, representative for our experience with these models. We first run the branch&cut algorithms as described in Sect. 7 from scratch. These results are summarized in Table 8.1. In a second set of runs we supplied optimum solutions as input to see how the implementations behave if the best upper bound is provided. The results are presented in Table 8.2.

---

Information on how to download the problem instances mentioned in this chapter can be obtained via http://www.zib.de/ascheuer/ATSTDWinstances.html
Solving the Asymmetric Travelling Salesman Problem with time windows by branch-and-cut

Table 8.1 Model comparison

<table>
<thead>
<tr>
<th>Model</th>
<th>Cuts</th>
<th>LP</th>
<th>PRIC</th>
<th>UB</th>
<th>HEUR</th>
<th>OPT</th>
<th>Cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>134</td>
<td>52</td>
<td>548</td>
<td>308</td>
<td>0.0379</td>
<td>214</td>
<td>206</td>
</tr>
<tr>
<td>Model 2</td>
<td>134</td>
<td>52</td>
<td>548</td>
<td>308</td>
<td>0.0379</td>
<td>214</td>
<td>206</td>
</tr>
</tbody>
</table>

Table 8.2 Model comparison (optimal solution supplied)

<table>
<thead>
<tr>
<th>Model</th>
<th>Cuts</th>
<th>LP</th>
<th>PRIC</th>
<th>UB</th>
<th>HEUR</th>
<th>OPT</th>
<th>Cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>134</td>
<td>52</td>
<td>548</td>
<td>308</td>
<td>0.0379</td>
<td>214</td>
<td>206</td>
</tr>
<tr>
<td>Model 2</td>
<td>134</td>
<td>52</td>
<td>548</td>
<td>308</td>
<td>0.0379</td>
<td>214</td>
<td>206</td>
</tr>
</tbody>
</table>

---

**Solving the Asymmetric Travelling Salesman Problem with time windows by branch-and-cut**

**Table 8.1 Model comparison**

<table>
<thead>
<tr>
<th>Model</th>
<th>Cuts</th>
<th>LP</th>
<th>PRIC</th>
<th>UB</th>
<th>HEUR</th>
<th>OPT</th>
<th>Cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>134</td>
<td>52</td>
<td>548</td>
<td>308</td>
<td>0.0379</td>
<td>214</td>
<td>206</td>
</tr>
<tr>
<td>Model 2</td>
<td>134</td>
<td>52</td>
<td>548</td>
<td>308</td>
<td>0.0379</td>
<td>214</td>
<td>206</td>
</tr>
</tbody>
</table>

**Table 8.2 Model comparison (optimal solution supplied)**

<table>
<thead>
<tr>
<th>Model</th>
<th>Cuts</th>
<th>LP</th>
<th>PRIC</th>
<th>UB</th>
<th>HEUR</th>
<th>OPT</th>
<th>Cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>134</td>
<td>52</td>
<td>548</td>
<td>308</td>
<td>0.0379</td>
<td>214</td>
<td>206</td>
</tr>
<tr>
<td>Model 2</td>
<td>134</td>
<td>52</td>
<td>548</td>
<td>308</td>
<td>0.0379</td>
<td>214</td>
<td>206</td>
</tr>
</tbody>
</table>
We infer from Tables 8.1 and 8.2 that Model 2 is the clear winner. The implementation of Model 2 ran fastest in 16 out of 22 cases, was second in 2 and third in 4 cases. Model 3 won 6 test runs, and was third in all others, while Model 1 came in second in all but two cases where it was third.

A more detailed analysis also shows that one should not expect one model to be the best over the whole range of applications. Model 2 did particularly well on problems of type \( \text{RBG}^a \) where time windows are active for all nodes (the usual case in our application). Our implementation of Model 3 works rather poorly in this problem range, but it outperforms the others when only a few time windows are active.

There are additional indicators that supported our choice of Model 2. We briefly mention a few of them. Our algorithm for Model 3 often ran into severe numerical problems. The linking constraints (3.6)(4) of this model caused problems whenever \( d_i = \infty \). Indeed, in case an LP variable \( x_{ij} \) has a very small value (slightly above the zero tolerance of the LP-solver) the corresponding \( y_{ij} \)–variables are not necessarily forced to 0. Therefore, adjustments with parameter settings of the LP solver are necessary. Moreover, the primal simplex method of the LP solver we used often returned an error code indicating that the problem was solved to optimality but *indeterminate infeasibilities* have been detected (CPLEX – infeasibility type 11). Finally, the resulting LPs of Model 3 are extremely difficult to solve and about 90% of the total computing time was spent within the LP solver.

The LP formulations of all three models differ significantly in their number of variables, number of constraints and their degree of difficulty. The LPs resulting from Model 2 seem to be the smallest and easiest for the LP solver. The numerical problems for Model 1 (big \( M \)) and Model 3 (linking constraints) result in high computing times to solve the LPs. Moreover, an optimal solution of the LP relaxation of Models 1 and 3 tends to have more fractional components, i.e., variables \( x^*_{ij} \) with \( 0 < x^*_{ij} < 1 \).

---

**Key to Table 8.3:**

<table>
<thead>
<tr>
<th>( n )</th>
<th>Number of nodes. There are ( n - 2 ) “real” nodes. For implementation reasons, the depot node has been split into a starting depot with no arc entering this node and into a terminal depot with no arc leaving this node.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(</td>
<td>A</td>
</tr>
<tr>
<td><strong>SOLUTION ... OPT</strong></td>
<td></td>
</tr>
<tr>
<td><strong>GAP</strong></td>
<td>Optimality gap in percent; calculated by ( \frac{gub - glb}{glb} \cdot 100 ).</td>
</tr>
<tr>
<td><strong>ROOT ... BOUNDS</strong></td>
<td>Lower bound ( rlb ) and upper bound ( rub ) at the root LP.</td>
</tr>
<tr>
<td><strong>GAP</strong></td>
<td>Optimality gap at the root of the BC-tree in percent; calculated by ( \frac{rub - rlb}{rlb} \cdot 100 ).</td>
</tr>
<tr>
<td><strong>QUAL</strong></td>
<td>Quality of lower bound at the root of the BC-tree in percent; calculated by ( 100 - \frac{(rub - rlb)}{rub} \cdot 100 ).</td>
</tr>
<tr>
<td><strong>BC-TREE ... #N</strong></td>
<td>Number of nodes in the branch&amp;cut tree.</td>
</tr>
<tr>
<td><strong>LEV</strong></td>
<td>Depth of the branch&amp;cut tree.</td>
</tr>
<tr>
<td><strong>#CUTS</strong></td>
<td>Number of generated cutting planes.</td>
</tr>
<tr>
<td><strong>#LP</strong></td>
<td>Number of linear programs that had to be solved.</td>
</tr>
<tr>
<td><strong>CPU</strong></td>
<td>Total CPU-time in minutes to solve problem instance to optimality. If the problem instance could not be solved to optimality within a certain time limit, this is marked by giving the CPU-time after which the run is stopped.</td>
</tr>
</tbody>
</table>
Solving the Asymmetric Travelling Salesman Problem with time windows by branch-and-cut

### Key to Table 8.4:
- **POOL**: Number of inequalities generated from the pool / Number of calls of pool separation.
- **SEC**: Number of generated subtour elimination constraints (SEC) / Number of calls of SEC separation routine.
- **TMC**: Number of generated 2–matching constraints / Number of calls of 2–matching separation routine.
- **π**: Number of generated π inequalities / Number of calls of π separation routine.
- **σ**: Number of generated σ inequalities / Number of calls of σ separation routine.
- **(π, σ)**: Number of generated (π, σ) inequalities / Number of calls of (π, σ) separation routine.
- **IPEC**: Number of generated infeasible path elimination constraints (IPEC) / Number of calls of IPEC separation routine.
- **SHRINK**: Number of inequalities generated by shrinking procedure (see Sect. 7.2) / Number of calls.
- **D_k**: Number of generated D_k–inequalities / Number of calls of D_k separation routine.
- **T_k**: Number of generated T_k–inequalities / Number of calls of T_k separation routine.
- **SD**: Number of generated SD–inequalities / Number of calls of SD separation routine.

### Key to Table 8.5:
- **TOURN**: Number of generated tournament constraints (4.1).
- **CONCAT**: Number of generated concatenated infeasible path elimination constraints (4.7).
- **IP1A**: Number of generated infeasible path elimination constraints (IPEC) (4.2) by enumeration procedure.
- **IP1B**: Number of generated IPEC (4.2) by modification of already detected infeasible paths.
- **IP2A**: Number of generated IPEC (4.3) by enumeration procedure.
- **IP2B**: Number of generated IPEC (4.3) by modification of already detected infeasible paths.
- **IP3A**: Number of generated IPEC (4.4) by enumeration procedure.
- **IP3B**: Number of generated IPEC (4.4) by modification of already detected infeasible paths.
- **IP4A**: Number of generated IPEC (4.5) by enumeration procedure.
- **IP4B**: Number of generated IPEC (4.5) by modification of already detected infeasible paths.
- **TOTAL**: Total number of generated infeasible path elimination constraints.

### Key to Table 8.6:
- **INIT**: Computing time spent in initialization phase (in %).
- **LP**: Computing time spent in LP-solver (in %).
- **IMPROVE**: Computing time spent in LP-exploitation and subsequent improvement heuristics (in %).
- **SEPARATION**: Computing time spent in various separation procedures (in %).
- **PRICING**: Computing time spent in pricing (in %).
- **MISC**: Computing time spent in other parts of the program (in %).
- **TOTAL**: Total CPU-time in minutes.

In Tables 8.1 and 8.2 we report computational results for instances up to 69 nodes only, as only instances up to this size could be solved to optimality by the implementations of all three models. Larger instances could only be solved by Model 2.

For all these reasons we decided to work with Model 2 and develop a more advanced implementation based on this model.

### 8.2. Advanced implementation of Model 2

The advanced implementation of the branch&cut-algorithm based on Model 2 was done using a preliminary version of the general purpose branch&cut-framework ABACUS that is explained in [29], the LP-solver CPLEX 5.0, and the FORTRAN implementations of the separation routines for SD- and D_k–inequalities by Fischetti and Toth [23].

We demonstrate here the performance of this code on a testbed of 50 problem instances varying from 12 to 233 nodes that were derived from a practical application.
Table 8.3: Computational results for the advanced implementation of Model 2

<table>
<thead>
<tr>
<th>INSTANCE</th>
<th>AVERAGE TIME</th>
<th>CPU TIME</th>
<th>LB</th>
<th>UB</th>
<th>GAP</th>
<th>OPT</th>
<th>LEV</th>
<th>ROOT</th>
<th>CPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>rbg041a</td>
<td>21304</td>
<td>10482</td>
<td>45</td>
<td>45</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>25184</td>
<td>328</td>
</tr>
<tr>
<td>rbg035a</td>
<td>1084</td>
<td>1111</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>2396</td>
<td>29</td>
</tr>
<tr>
<td>rbg031a</td>
<td>1797</td>
<td>1400</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>23396</td>
<td>300</td>
</tr>
<tr>
<td>rbg019a</td>
<td>12809</td>
<td>13204</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>20294</td>
<td>58</td>
</tr>
<tr>
<td>rbg017a</td>
<td>1523</td>
<td>13204</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>23396</td>
<td>300</td>
</tr>
<tr>
<td>rbg015a</td>
<td>16009</td>
<td>13204</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>20294</td>
<td>58</td>
</tr>
<tr>
<td>rbg013a</td>
<td>21304</td>
<td>10482</td>
<td>45</td>
<td>45</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>25184</td>
<td>328</td>
</tr>
<tr>
<td>rbg011a</td>
<td>1084</td>
<td>1111</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>2396</td>
<td>29</td>
</tr>
<tr>
<td>rbg009a</td>
<td>1797</td>
<td>1400</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>23396</td>
<td>300</td>
</tr>
<tr>
<td>rbg007a</td>
<td>12809</td>
<td>13204</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>20294</td>
<td>58</td>
</tr>
<tr>
<td>rbg005a</td>
<td>1523</td>
<td>13204</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>23396</td>
<td>300</td>
</tr>
<tr>
<td>rbg003a</td>
<td>21304</td>
<td>10482</td>
<td>45</td>
<td>45</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>25184</td>
<td>328</td>
</tr>
<tr>
<td>rbg001a</td>
<td>1084</td>
<td>1111</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>100</td>
<td>17</td>
<td>2396</td>
<td>29</td>
</tr>
</tbody>
</table>

For all instances we allow a maximum computation time of 5 hours of CPU-time. All runs are executed with the default parameter settings described in the previous section. If an instance could not be solved to optimality within the given time limit, we list the lower and upper bounds found by the algorithm. The results are summarized in Tables 8.3–8.6. Table 8.3 gives general information on the performance of the implementation, Table 8.4 (resp. Table 8.5) on the number of generated cutting planes (resp. generated infeasible path elimination constraints). Table 8.6 summarizes the percentages of computing time spent in the various parts of the program.
Solving the Asymmetric Travelling Salesman Problem with time windows by branch-and-cut

The instances of our application, either optimal solutions or feasible solutions within an acceptable quality guarantee, was solved to optimality within 6 seconds. Only 176 cutting planes were generated for the instances of our application, either optimal solutions or feasible solutions within an acceptable quality guarantee.
From Table 8.3 one can derive that the heuristics embedded in our implementation found in 21 out of 50 cases an optimal solution in the root node of the branch&cut-tree. For these instances the optimality of the solutions was proven fast. For all instances that were solved to optimality the GAP between the lower and upper bound before branching is relatively small. For all such instances, except rgb035a.2 and rgb040a it is within a 7% range; for most of the instances even smaller (see column \textit{ROOT} . . . \textit{GAP}).

For the instances that were not solved to optimality only minor improvements in the upper bound could be obtained in the branching phase, whereas the lower bound could be improved significantly (cmp. columns \textit{SOLUTION} . . . \textit{OPT} and \textit{ROOT} . . . \textit{BOUNDS}).

The largest instance (in terms of nodes) solved to optimality has 127 nodes. But one should note as well that there exist small instances (e.g., rgb019b) that were hard to solve by the branch&cut-algorithm. The smallest unsolved instance contains 43 nodes.
Table 8.6 Percentage of computing time spent in different parts of algorithm

<table>
<thead>
<tr>
<th></th>
<th>INT</th>
<th>LP</th>
<th>IMPROVE</th>
<th>SEPARATION</th>
<th>PRICING</th>
<th>MISC</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>rgb010a</td>
<td>0.0</td>
<td>14.3</td>
<td>14.3</td>
<td>28.6</td>
<td>0.0</td>
<td>42.9</td>
<td>0.00.12</td>
</tr>
<tr>
<td>rgb016a</td>
<td>16.7</td>
<td>33.3</td>
<td>8.3</td>
<td>41.7</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00.20</td>
</tr>
<tr>
<td>rgb016b</td>
<td>0.4</td>
<td>32.6</td>
<td>13.1</td>
<td>36.0</td>
<td>4.9</td>
<td>13.1</td>
<td>0.00.80</td>
</tr>
<tr>
<td>rgb017.2</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00.03</td>
</tr>
<tr>
<td>rgb017</td>
<td>4.1</td>
<td>24.5</td>
<td>16.3</td>
<td>44.9</td>
<td>0.0</td>
<td>10.2</td>
<td>0.00.82</td>
</tr>
<tr>
<td>rgb017a</td>
<td>57.1</td>
<td>42.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00.12</td>
</tr>
<tr>
<td>rgb019a</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00.03</td>
</tr>
<tr>
<td>rgb019b</td>
<td>0.1</td>
<td>23.5</td>
<td>11.4</td>
<td>40.2</td>
<td>4.6</td>
<td>20.3</td>
<td>0.54.57</td>
</tr>
<tr>
<td>rgb019c</td>
<td>1.1</td>
<td>32.5</td>
<td>7.5</td>
<td>42.6</td>
<td>2.7</td>
<td>13.6</td>
<td>0.08.72</td>
</tr>
<tr>
<td>rgb019d</td>
<td>2.2</td>
<td>26.7</td>
<td>13.3</td>
<td>48.9</td>
<td>0.0</td>
<td>8.9</td>
<td>0.00.75</td>
</tr>
<tr>
<td>rgb020a</td>
<td>91.7</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>8.3</td>
<td>0.0</td>
<td>0.00.20</td>
</tr>
<tr>
<td>rgb021.2</td>
<td>38.5</td>
<td>38.5</td>
<td>0.0</td>
<td>15.4</td>
<td>7.7</td>
<td>0.0</td>
<td>0.00.22</td>
</tr>
<tr>
<td>rgb021.3</td>
<td>0.3</td>
<td>25.8</td>
<td>7.1</td>
<td>50.0</td>
<td>3.3</td>
<td>13.6</td>
<td>0.27.15</td>
</tr>
<tr>
<td>rgb021.4</td>
<td>1.4</td>
<td>25.2</td>
<td>4.9</td>
<td>44.7</td>
<td>3.2</td>
<td>20.6</td>
<td>0.05.82</td>
</tr>
<tr>
<td>rgb021.5</td>
<td>1.3</td>
<td>27.9</td>
<td>5.8</td>
<td>37.9</td>
<td>5.8</td>
<td>21.4</td>
<td>0.06.63</td>
</tr>
<tr>
<td>rgb021.6</td>
<td>7.2</td>
<td>44.6</td>
<td>18.1</td>
<td>22.9</td>
<td>3.6</td>
<td>3.6</td>
<td>0.01.38</td>
</tr>
<tr>
<td>rgb021.7</td>
<td>1.9</td>
<td>26.0</td>
<td>27.9</td>
<td>26.7</td>
<td>1.9</td>
<td>15.5</td>
<td>0.04.30</td>
</tr>
<tr>
<td>rgb021.8</td>
<td>0.5</td>
<td>26.2</td>
<td>24.1</td>
<td>24.3</td>
<td>3.9</td>
<td>20.9</td>
<td>0.17.40</td>
</tr>
<tr>
<td>rgb021.9</td>
<td>0.4</td>
<td>23.5</td>
<td>26.9</td>
<td>22.5</td>
<td>5.7</td>
<td>21.0</td>
<td>0.26.12</td>
</tr>
<tr>
<td>rgb021</td>
<td>1.0</td>
<td>34.9</td>
<td>5.3</td>
<td>44.6</td>
<td>3.0</td>
<td>11.2</td>
<td>0.08.75</td>
</tr>
<tr>
<td>rgb027a</td>
<td>14.5</td>
<td>28.1</td>
<td>7.4</td>
<td>36.3</td>
<td>5.9</td>
<td>7.4</td>
<td>0.02.25</td>
</tr>
<tr>
<td>rgb031a</td>
<td>12.7</td>
<td>24.5</td>
<td>31.4</td>
<td>25.5</td>
<td>2.0</td>
<td>3.9</td>
<td>0.01.70</td>
</tr>
<tr>
<td>rgb033a</td>
<td>16.2</td>
<td>18.9</td>
<td>21.6</td>
<td>29.7</td>
<td>0.0</td>
<td>13.5</td>
<td>0.01.85</td>
</tr>
<tr>
<td>rgb034a</td>
<td>40.7</td>
<td>25.4</td>
<td>8.5</td>
<td>8.5</td>
<td>11.9</td>
<td>5.1</td>
<td>0.00.98</td>
</tr>
<tr>
<td>rgb035a.2</td>
<td>1.0</td>
<td>33.8</td>
<td>23.7</td>
<td>19.1</td>
<td>10.1</td>
<td>12.3</td>
<td>1.04.80</td>
</tr>
<tr>
<td>rgb035a</td>
<td>19.1</td>
<td>24.5</td>
<td>28.2</td>
<td>21.8</td>
<td>0.9</td>
<td>5.5</td>
<td>0.01.83</td>
</tr>
<tr>
<td>rgb038a</td>
<td>0.0</td>
<td>13.8</td>
<td>5.8</td>
<td>61.2</td>
<td>2.3</td>
<td>16.9</td>
<td>7.03.23</td>
</tr>
<tr>
<td>rgb040a</td>
<td>0.1</td>
<td>23.1</td>
<td>5.1</td>
<td>49.8</td>
<td>4.5</td>
<td>17.4</td>
<td>12.31.82</td>
</tr>
<tr>
<td>rgb041a</td>
<td>0.0</td>
<td>25.2</td>
<td>10.6</td>
<td>54.5</td>
<td>2.2</td>
<td>7.5</td>
<td>300.01.68</td>
</tr>
<tr>
<td>rgb042a</td>
<td>0.0</td>
<td>22.1</td>
<td>16.5</td>
<td>51.3</td>
<td>2.5</td>
<td>7.6</td>
<td>300.01.68</td>
</tr>
<tr>
<td>rgb048a</td>
<td>0.0</td>
<td>20.2</td>
<td>3.1</td>
<td>62.1</td>
<td>5.1</td>
<td>9.5</td>
<td>300.01.12</td>
</tr>
<tr>
<td>rgb049a</td>
<td>0.0</td>
<td>16.0</td>
<td>3.5</td>
<td>69.8</td>
<td>3.6</td>
<td>7.1</td>
<td>300.01.22</td>
</tr>
<tr>
<td>rgb050a</td>
<td>7.2</td>
<td>16.2</td>
<td>44.2</td>
<td>19.6</td>
<td>3.7</td>
<td>12.0</td>
<td>0.18.62</td>
</tr>
<tr>
<td>rgb050b</td>
<td>0.0</td>
<td>21.0</td>
<td>2.1</td>
<td>61.5</td>
<td>9.6</td>
<td>5.8</td>
<td>300.01.33</td>
</tr>
<tr>
<td>rgb050c</td>
<td>0.0</td>
<td>18.9</td>
<td>5.3</td>
<td>60.2</td>
<td>4.5</td>
<td>11.1</td>
<td>300.01.27</td>
</tr>
<tr>
<td>rgb055c</td>
<td>22.2</td>
<td>20.6</td>
<td>16.9</td>
<td>26.6</td>
<td>2.6</td>
<td>11.2</td>
<td>0.06.40</td>
</tr>
<tr>
<td>rgb067a</td>
<td>41.2</td>
<td>21.0</td>
<td>17.4</td>
<td>14.3</td>
<td>1.4</td>
<td>4.8</td>
<td>0.05.95</td>
</tr>
<tr>
<td>rgb086a</td>
<td>0.0</td>
<td>6.6</td>
<td>21.1</td>
<td>39.3</td>
<td>2.5</td>
<td>30.5</td>
<td>0.01.95</td>
</tr>
<tr>
<td>rgb092a</td>
<td>0.0</td>
<td>16.0</td>
<td>9.0</td>
<td>49.2</td>
<td>5.3</td>
<td>20.5</td>
<td>300.01.40</td>
</tr>
<tr>
<td>rgb125a</td>
<td>7.8</td>
<td>8.3</td>
<td>17.1</td>
<td>16.4</td>
<td>3.9</td>
<td>46.4</td>
<td>3.49.82</td>
</tr>
<tr>
<td>rgb132.2</td>
<td>0.2</td>
<td>13.6</td>
<td>9.4</td>
<td>55.1</td>
<td>4.8</td>
<td>16.9</td>
<td>300.01.42</td>
</tr>
<tr>
<td>rgb132</td>
<td>0.1</td>
<td>13.0</td>
<td>8.9</td>
<td>53.5</td>
<td>3.6</td>
<td>20.9</td>
<td>300.01.47</td>
</tr>
<tr>
<td>rgb152s</td>
<td>0.5</td>
<td>42.8</td>
<td>16.9</td>
<td>16.1</td>
<td>9.8</td>
<td>14.0</td>
<td>300.02.90</td>
</tr>
<tr>
<td>rgb152</td>
<td>0.2</td>
<td>11.4</td>
<td>12.6</td>
<td>44.0</td>
<td>4.1</td>
<td>27.7</td>
<td>300.03.10</td>
</tr>
<tr>
<td>rgb172a</td>
<td>0.3</td>
<td>11.5</td>
<td>7.2</td>
<td>59.7</td>
<td>5.1</td>
<td>16.1</td>
<td>300.02.43</td>
</tr>
<tr>
<td>rgb193c</td>
<td>0.7</td>
<td>29.7</td>
<td>9.2</td>
<td>30.2</td>
<td>8.9</td>
<td>21.3</td>
<td>300.06.55</td>
</tr>
<tr>
<td>rgb193</td>
<td>0.4</td>
<td>8.3</td>
<td>6.9</td>
<td>52.0</td>
<td>7.8</td>
<td>24.6</td>
<td>300.03.98</td>
</tr>
<tr>
<td>rgb201a</td>
<td>0.5</td>
<td>7.8</td>
<td>6.1</td>
<td>56.8</td>
<td>7.9</td>
<td>20.8</td>
<td>300.06.08</td>
</tr>
<tr>
<td>rgb233.2</td>
<td>1.4</td>
<td>28.5</td>
<td>17.4</td>
<td>23.2</td>
<td>8.1</td>
<td>21.4</td>
<td>300.01.28</td>
</tr>
<tr>
<td>rgb233</td>
<td>0.7</td>
<td>13.6</td>
<td>11.5</td>
<td>55.3</td>
<td>3.8</td>
<td>14.6</td>
<td>300.07.62</td>
</tr>
</tbody>
</table>

Table 8.4 gives, for each separation routine, the number of generated cuts in relation to the number of times this separation routine was called. For example, on instance rgb016b the separation routine for the \(\pi\)–inequalities was called 148 times generating 41 cutting planes in total. One sees that very seldomly violated 2–matching constraints, \(D_{\kappa}\)–, and \(T_{\kappa}\)–inequalities were found. Among the other separation routines the number of generated cuts is more or less evenly distributed.

Table 8.5 gives a more detailed view on the number of infeasible path elimination constraints (IPEC) that were generated. Most of the generated IPEC are tournament constraints. On some instances (e.g., rgb041a) most of the IPEC are tournament con-
straits or inequalities of type (4.5). Only on very few instances a large number of violated inequalities (4.2)–(4.4) were found. Surprisingly, violated inequalities of type (4.4) and (4.5) were never found by the separation heuristic checking modifications of already found infeasible paths.

9. Conclusions

Our computational experience indicates that most ATSP-TW instances in the range of up to 50–70 nodes can be solved to optimality via branch&cut codes based on any of the three models. The LPs arising from Models 1 and 3 are larger and, at times, significant numerical instabilities occur in the solution process, despite the fact that our LP solver CPLEX is generally very good at handling numerical difficulties. The reasons for this are model inherent (big $M$, weak linking constraints) and not due to poor software. Overall the implementation of Model 2 outperformed the two other codes, although for special cases (few time windows active) Model 3 appeared to be more suitable.

The results of our test runs made us conclude that Model 2 is suited best for our particular application of ATSP-TW and we decided to produce a more thorough implementation. This branch&cut code (based on ABACUS and CPLEX 5.0) was tested on real–world and modified real–world instances with sizes of up to about 250 nodes. However, this model has the disadvantage that it can only handle TSP–objective functions but not makespan objectives and the like. For our application, this does not matter, though.

We should also remark here that our computational experience does not indicate that – in contrast to the TSP, say – the polyhedral approach to the ATSP-TW is the unchallenged winning strategy. We believe that dynamic programming and implicit enumeration techniques may outperform our cutting plane method, in particular when the time windows are rather tight. Further research (including very time consuming implementational work) has to show where the relative advantages of the different methodologies lie. Nevertheless, the heuristics and cutting plane algorithms described in this paper were able to solve real–world instances of a particular application (stacker crane optimization) of the asymmetric travelling salesman problem with time windows in a way that is satisfactory for practice.

References


² Available at URL http://www.zib.de/ZIBbib/Publications/


37. Savelsbergh, M. (1994): Personal communication. School of Industrial and System Engineering, Georgia Institute of Technology, Atlanta, USA


40. Simonetti, N. (1997): Personal communication


