Data for research on nurse rostering in Denmark

Elín Björk Böðvarsdóttir
Department of Technology, Management and Economics, Technical University of Denmark

Niels-Christian Fink Bagger
Fink Optimization

Laura Elise Høffner
Department of Data and Development Support, Region Zealand

Thomas Stidsen
Department of Technology, Management and Economics, Technical University of Denmark

May 2019

1 Introduction

This report describes data for nurse rostering in the Danish setting, and is supplementary for the technical report *A comprehensive integer programming formulation of the nurse rostering problem in Denmark*. We present 12 instances for 2 wards in 2 Danish hospitals in Region Zealand.

The structure of this report is as follows: Section 2 presents the structure of the data and Section 3 provides deeper insight into the data, to ensure the reproducibility of the research.

2 Structure

Each instance consists of 21 `.csv` files, where the parameters have been grouped together based on their indices. If a file is not included for an instance then the corresponding constraints are inactive. We note that although shift types and shift sets are often interchangeable, we present all the data on a shift set level for simplicity. The files for each instance are as follows:

- `General.csv`
- `Weeks.csv`
- `ShiftTypes.csv`
- `ShiftSets.csv`
- `ShiftTypes_ShiftSets.csv`
- `Coverage.csv`
- `Coverage_Date.csv`
- `Employees.csv`
- `FeasibleAssignments.csv`
- `PreviousHorizon.csv`
- `EmployeesTogether.csv`
- `Employee_Week.csv`
- `Employee_ShiftSet.csv`

- `Employee_ShiftSet_Week.csv`
- `Employee_ShiftSet_Date.csv`
- `SurroundedDaysOff.csv`
- `Patterns.csv`
- `Employee_Date_Pattern.csv`
- `AssignmentLeadsToOther_Ward.csv`
- `AssignmentLeadsToOTher_Employee.csv`
- `Requests.csv`

In the following subsections we will present the contents of each file and relate them to the notation used in the article.

2.1 General

This file contains two columns. The former column gives the parameter name and the latter column gives the parameter value, given in the format `Parameter;Value`. The entries are as follows:

- `FirstDateCurrent`: The first date of the planning horizon.
- `LastDate`: The last date of the planning horizon.
  - The days of the planning horizon $\mathcal{D}$ include `FirstDateCurrent`, `LastDate` and all the days in between.
- `FirstDatePrevious`: The earliest date we consider from previous plans.
  - The days of previous plans $\mathcal{D}^{pre}$ include `FirstDatePrevious` and all days up to (but not including) `FirstDateCurrent`.
  - All days we consider is the union $\mathcal{D}^{all} = \mathcal{D} \cup \mathcal{D}^{pre}$.
- `rho^{full}`: The minimum time ρ^{full} we require for full rest between assignments.
- `rho^{red}`: The minimum time ρ^{red} we require for reduced rest between assignments.
  - We only include this line if the corresponding ward allows reduced rest, and then the value is strictly lower than the full rest ρ^{full}.
  - We use the requirements for rest (along with start and end times for shift types) in constraint H2. More precisely, we generate the conflict cliques Γ, conflicting assignments $\mathcal{A}_\gamma$ for $\gamma \in \Gamma$ and work conflict cliques $\Gamma_{e,d}$ for employee $e \in \mathcal{E}$ on day $d \in \mathcal{D}$ as described in Section 3.1.1 in the article. However, we have two exceptions where shift types overlapping in time are not considered as conflicting. Those exceptions are assigning either `FO` or `PF` on a day following a night shift, where we refer to Section 3.1 for a discussion on the naming conventions of shift types.
- `T^{prev}`: A reference time T^{prev} from previous horizon.
- `T^{fut}`: A reference time T^{fut} from future horizon.
- `rho^{pf}_{i}`: The minimum hours off ρ_i^{pf} for $i \in \mathbb{N}$ consecutive protected days off.
  - We provide this entry in multiple lines for $i \in \{1, \ldots, n\}$ where n is the total number of protected days off that we should assign during the planning horizon.
  - We use the reference times and minimum hours off in constraint H3.
- `v^{pref}`: The maximum preferred distance v^{pref} between two contiguous protected days off, according to constraint S2.
- `N^{+}`: The maximum positive deviation N^+ from target hours before penalizing heavier in two-factor penalty.

- `N^{-}`: The maximum negative deviation N^- from target hours before penalizing heavier in two-factor penalty.
  - We use the bounds on deviation from target hours in constraint S3.
- `n^{diff}`: The maximum number n^{diff} of different shift sets from $\mathcal{Z}^{diff}$ we prefer on consecutive days.
- `D^{diff}`: The number of consecutive days D^{diff} when restricting the number of different shift sets.
  - We address the number of different shift sets on consecutive days with constraints S8.
- `m^{seq}`: The minimum preferred length m^{seq} of working sequences, according to constraint S9.
- `l^{gen}`: The standard length l^{gen} of a shift type, which we use for constraint S15. If the constraint is inactive then we exclude this line from the data.
- `omega^{pf}`: The weight ω^{pf} in constraint S2.
- `omega^{+}`: The weight ω^{+} in constraint S3. If $N^+ = 0$ then we exclude this line from the data.
- `omega^{++}`: The weight ω^{++} in constraint S3.
- `omega^{-}`: The weight ω^{-} in constraint S3. If $N^- = 0$ then we exclude this line from the data.
- `omega^{-}`: The weight ω^{--} in constraint S3.
- `omega^{week}`: The weight ω^{week} in constraint S4.
- `omega^{diff}`: The weight ω^{diff} in constraint S7.
- `omega^{seqlen}`: The weight ω^{seqlen} in constraint S8.
- `omega^{seqconsec}`: The weight $\omega^{seqconsec}$ in constraint S9. If the constraint is inactive then we exclude this line from the data.
- `omega^{b}`: The weight ω^{b} in constraint S10. If the constraint is inactive then we exclude this line from the data.
- `omega^{a}`: The weight ω^{a} in constraint S10. If the constraint is inactive then we exclude this line from the data.
- `omega^{bal}`: The weight ω^{bal} in constraint S13.

2.2 Weeks

This file contains 2 columns, as presented with Listing 1. The former is an integer denoting the identity of the week and the latter is a set containing all the days associated with the week. These entries define the sets $\mathcal{W}$ for weeks of the planning horizon and $\mathcal{D}_w$ of the days in week $w \in \mathcal{W}$.

```
1 WeekId;Dates
2 1;{'2018-09-10', '2018-09-13', '2018-09-12', '2018-09-14', '2018-09-11', '2018-09-15',
      '2018-09-16'}
3 2;{'2018-09-22', '2018-09-17', '2018-09-20', '2018-09-21', '2018-09-23', '2018-09-19',
      '2018-09-18'}
```

Listing 1: `Weeks.csv` format.

2.3 ShiftTypes

This file contains 4 columns, as presented with Listing 2. All entries in this file are shift types $s \in \mathcal{S}$. The columns are as follows:

- `ShiftId`: The id of the shift type.
- `WorkShift`: A boolean indicating whether the shift type is a work shift, used to define the set $\mathcal{S}^{work}$.
- `StartTime`: The starting time of the shift type.
- `Duration`: The duration of the shift type.
  - From the starting time and duration of a shift type we can calculate the parameters $T_{s,d}^{start}$ and $T_{s,d}^{end}$ for every day $d \in \mathcal{D}^{all}$. We use these parameters in various ways, e.g., when evaluating conflicting assignments and protected days off.

In Section 3.1 we discuss the identities we use for different shift types.

```
1 ShiftId;WorkShift;StartTime;Duration
2 PF;False;00:00:00;24.0
3 D11;True;07:00:00;8.0
4 MENT;True;07:00:00;8.25
5 D66;True;07:00:00;12.25
```

Listing 2: `ShiftTypes.csv` format.

2.4 ShiftSets

This file contains up to 5 columns, as presented with Listing 3. All entries in this file are shift sets $\sigma \in \mathcal{Z}$. The columns are as follows:

- `SetId`: The id of the shift set.
- `In Z^{diff}`: A boolean indicating whether the shift set is in $\mathcal{Z}^{diff}$.
- `In Z^{together}`: A boolean indicating whether the shift set is in $\mathcal{Z}^{together}$. If constraint S15 is inactive then we exclude this column from the data.
- `In Z^{spread}`: A boolean indicating whether the shift set is in $\mathcal{Z}^{spread}$. If constraint S14 is inactive then we exclude this column from the data.
- `omega^{spread}_{sigma}`: The weight ω_{σ}^{spread} in constraint S14. If the constraint is inactive then we exclude this column from the data.

```
1 SetId;In Z^{diff};In Z^{together};In Z^{spread};omega^{spread}_{sigma}
2 A;False;True;False;
3 AorN;False;False;True;49950.0
4 D;False;False;False;
```

Listing 3: `ShiftSets.csv` format.

2.5 ShiftTypes_ShiftSets

This file contains 2 columns, as presented with Listing 4. The former column is a tuple (s, σ) for $\sigma \in \mathcal{Z}$ and $s \in \mathcal{S} \cap \sigma$, and the latter column is the relation $\alpha_{s,\sigma}$ between the shift type and the shift set.

```
1 (s,sigma);alpha_{s,sigma}
2 (A92, AorN);1.5
3 (N56, AorN);1.0
4 (D85, Evening);0.5
5 (A02, Evening);1.0
6 (A36, Evening);0.5
7 (PF, Free);1.0
8 (A36, Night);1.0
```

Listing 4: `ShiftTypes_ShiftSets.csv` format.

2.6 Coverage

This file contains up to 6 columns, as presented with Listing 5. All entries in this file are coverage constraints $j \in \mathcal{C}$. The columns are as follows:

- `CoverageId`: The id of the coverage.
- `sigma_{j}`: The shift set $\sigma_j \in \mathcal{Z}$ corresponding to coverage $j \in \mathcal{C}$.
- `Xi_{j}`: The set of positions Ξ_j that we can assign to coverage $j \in \mathcal{C}$.
- `beta^{float}_{j}`: A boolean β_j^{float} denoting whetherwe can assign substitutes to coverage $j \in \mathcal{C}$.
- `omega^{float}_{j}`: The weight ω_j^{float} in constraint S11.
- `omega^{nonmax}_{j}`: The weight ω_j^{nonmax} in constraint S12. If the constraint is inactive then we exclude this column from the data.

```
1 CoverageId;sigma_{j};Xi_{j};beta^{float}_{j};omega^{float}_{j};omega^{nonmax}_{j}
2 1;Day;{'SHA', 'SpecialTrainedNurse', 'Nurse'};1;183750.0;
3 3;Evening;{'SHA', 'SpecialTrainedNurse', 'Nurse'};1;123750.0;50000.0
4 4;Evening;{'SpecialTrainedNurse'};1;123750.0;
```

Listing 5: `Coverage.csv` format.

2.7 Coverage_Date

This file contains 3 columns, as presented with Listing 6. The columns are as follows:

- `(j,d)`: A tuple denoting coverage $j \in \mathcal{C}$ and day $d \in \mathcal{D}$.
- `c^{min}_{j,d}`: The minimum number $c_{j,d}^{min}$ of employees we require for coverage $j \in \mathcal{C}$ on day $d \in \mathcal{D}$, used in constraint H11.
- `c^{max}_{j,d}`: The maximum number $c_{j,d}^{max}$ of employees we allow for coverage $j \in \mathcal{C}$ on day $d \in \mathcal{D}$, used in constraint H12.

```
1 (j,d);c^{min}_{j,d};c^{max}_{j,d}
2 (1, '2019-02-01');10;
3 (3, '2019-01-29');7;7
```

Listing 6: `Coverage_Date.csv` format.

2.8 Employees

This file contains up to 7 columns, as presented with Listing 7. All entries in this file are employees $e \in \mathcal{E}^{all}$. The columns are as follows:

- `EmployeeId`: The id of the employee.
- `CreateFullSchedule`: A boolean indicating whether the employee is in the set $\mathcal{E}$.
- `T^{target}_{e}`: The number of hours T_e^{target} that we should assign employee $e \in \mathcal{E}$.
- `xi_{e}`: The position ξ_e of employee $e \in \mathcal{E}^{all}$.
- `omega^{maxrow}_{e}`: The weight ω_e^{maxrow} in constraint S5.
- `IsChaperone`: A boolean indicating whether the employee is in the set $\mathcal{E}^{chap}$.
- `HasChaperone`: If the employee has a chaperone, then this column indicates the who that mentor is and thereby defines the set $\mathcal{E}_c^{train}$.

– If the ward does not use mentoring then we exclude the last two columns from the data.

```
EmployeeId;CreateFullSchedule;T^{target}_{e};xi_{e};omega^{maxrow}_{e};IsChaperone;HasChaperone
e00;True;124.85;SpecialTrainedNurse;1902.0;False;
e18;True;145.1;Nurse;1359.0;False;e25
e36;True;149.7;Nurse;1236.0;False;e09
e39;False;;SpecialTrainedNurse;;;
```

Listing 7: `Employees.csv` format.

2.9 FeasibleAssignments

This file contains 3 columns, as presented with Listing 8. All entries in this file are feasible assignments $(e, d, s) \in \mathcal{A}$. From the feasible assignments we can derive the set $\mathcal{S}_{e,d}$ of feasible shift types for employee $e \in \mathcal{E}^{all}$ on day $d \in \mathcal{D}$. The columns are as follows:

- `(e,d,s)`: A tuple denoting employee $e \in \mathcal{E}^{all}$, day $d \in \mathcal{D}$ and shift type $s \in \mathcal{S}$.
- `T^{hours}_{e,d,s}`: The hours $T^{hours}_{e,d,s}$ that assignment (e, d, s) counts towards the contractual hours.
- `omega^{assign}_{e,d,s}`: The weight $\omega^{assign}_{e,d,s}$ in constraint S1.

```
(e,d,s);T^{hours}_{e,d,s};omega^{assign}_{e,d,s}
(e00, '2018-09-12', PF);0.0;0.0
(e00, '2018-09-12', D49);8.25;-5833.0
(e03, '2018-09-13', N99);8.25;1389.0
(e12, '2018-09-20', FE);6.0;0.0
(e17, '2018-09-10', BD);7.0;0.0
```

Listing 8: `FeasibleAssignments.csv` format.

2.10 PreviousHorizon

This file contains one column, as presented with Listing 9. All entries are tuples (e, d, s) for $e \in \mathcal{E}, d \in \mathcal{D}^{pre}, s \in \mathcal{S}$ such that $x^{pre}_{e,d,s} = 1$.

```
x^{pre}_{e,d,s}=1
(e00, '2018-09-03', D49)
(e00, '2018-09-04', D49)
```

Listing 9: `PreviousHorizon.csv` format.

2.11 EmployeesTogether

This file contains two columns, as presented with Listing 10. The former column is a tuple $(e_1, e_2) \in \mathcal{E}^G$ and the latter is the weight $\omega^{together}_{e_1,e_2}$ in constraint S15. If constraint S15 is inactive then we exclude this file from the data.

```
(e1,e2);omega^{together}_{e1,e2}
(e14, e25);-20
(e18, e10);-10
```

Listing 10: `EmployeesTogether.csv` format.

2.12 Employee_Week

This file contains two columns, as presented with Listing 11. The former column is a tuple (e, w) for employee $e \in \mathcal{E}$ and week $w \in \mathcal{W}$ and the latter is the weekly target hours $T^{week}_{e,w}$ for the employee in the week, according to constraint S4.

```
(e,w);T^{week}_{e,w}
(e00, 1);12.64
(e00, 2);29.14
```

Listing 11: `Employee_Week.csv` format.

2.13 Employee_ShiftSet

This file contains up to 5 columns, as presented with Listing 12. The columns are as follows:

- `(e,sigma)`: A tuple (e, σ) for employee $e \in \mathcal{E}$ and shift set $\sigma \in \mathcal{Z}$.
- `M^{total,ub}_{e,sigma}`: The maximum number of assignments $M^{total,ub}_{e,\sigma}$ for employee $e \in \mathcal{E}$ to shift set $\sigma \in \mathcal{Z}$ during the planning horizon, according to constraint H4. The extension, $\tilde{M}^{total,ub}_{e,\sigma}$ (used in constraint S15) is derived as described in Section 3.2.2 in the article.
- `M^{total,fix}_{e,sigma}`: The total number of assignments $M^{total,fix}_{e,\sigma}$ employee $e \in \mathcal{E}$ should have to shift set $\sigma \in \mathcal{Z}$ during the planning horizon, according to constraint H5.
- `chi_{e,sigma}`: The maximum preferred number of sequences $\chi_{e,\sigma}$ for the employee containing the shift set on consecutive days, according to constraint S9.
- `D^{seq}_{e,sigma}`: The number of consecutive days when restricting sequences.
  - The last two columns define the set $\mathcal{Z}^{seq}_e$. If constraint S9 is inactive then we exclude these columns from the data.

```
1 (e,sigma);M^{total,ub}_{e,sigma};M^{total,fix}_{e,sigma};chi_{e,sigma};D^{seq}_{e,sigma}
2 (e00, AorN);6.0;;;
3 (e00, N);;;1;7
4 (e00, PF);;8.0;;
```

Listing 12: `Employee_ShiftSet.csv` format.

2.14 Employee_ShiftSet_Week

This file contains 2 columns, as presented with Listing 13. The columns are as follows:

- `(e,sigma,w)`: A tuple (e, σ, w) for employee $e \in \mathcal{E}$, shift set $\sigma \in \mathcal{Z}$ and week $w \in \mathcal{W}$.
- `M^{week,fix}_{e,sigma,w}`: The total number of assignments $M^{week,fix}_{e,\sigma,w}$ employee $e \in \mathcal{E}$ should have to shift set $\sigma \in \mathcal{Z}$ during week $w \in \mathcal{W}$, according to constraint H6.

```
1 (e,sigma,w);M^{week,fix}_{e,sigma,w}
2 (e17, BD, 1);1
```

Listing 13: `Employee_ShiftSet_Week.csv` format.

2.15 Employee_ShiftSet_Date

This file contains up to 5 columns, as presented with Listing 14. The columns are as follows:

- `(e,d,sigma)`: A tuple (e, d, σ) for employee $e \in \mathcal{E}$, day $d \in \mathcal{D}$ and shift set $\sigma \in \mathcal{Z}$.
- `M^{row}_{e,d,sigma}`: The maximum number of assignments in a row $M^{row}_{e,d,sigma}$ for employee $e \in \mathcal{E}$ to shift set $\sigma \in \mathcal{Z}$ from day $d \in \mathcal{D}$, according to constraints H7 and S5.
- `beta^{row}_{e,d,sigma}`: A boolean $\beta^{row}_{e,d,\sigma}$ denoting whether the maximum assignments in a row should be a hard constraint (H7) or a soft constraint (S5).
- `H^{consec}_{e,d,sigma}`: The maximum number of hours $H^{consec}_{e,d,\sigma}$ we can assign employee $e \in \mathcal{E}$ to shift set $\sigma \in \mathcal{Z}$ on consecutive days from day $d \in \mathcal{D}$, according to constraint H8.
- `D^{consec}_{e,d,sigma}`: The number of consecutive days $D^{consec}_{e,d,\sigma}$ when restricting the hours.

```
1 (e,d,sigma);M^{row}_{e,d,sigma};beta^{row}_{e,d,sigma};H^{consec}_{e,d,sigma};D^{consec}_{e,d,
      sigma}
2 (e00, '2018-09-03', A);3;1;;
3 (e00, '2018-09-04', WorkHours);;;52.0;7
```

Listing 14: `Employee_ShiftSet_Date.csv` format.

2.16 SurroundedDaysOff

This file contains up to 5 columns, as presented with Listing 15. All entries in this file are a set of periods $\mathcal{Q}_e$ where we prefer to assign surrounding days off for employee $e \in \mathcal{E}$, corresponding to constraint S10. If the constraint is inactive then we exclude this file from the data. The columns are as follows:

- `(e,q)`: A tuple (e, q) for employee $e \in \mathcal{E}$ and $q \in \mathcal{Q}_e$.
- `q^{start}`: The first day q^{start} of the period $q \in \mathcal{Q}_e$ for employee $e \in \mathcal{E}$.
- `q^{end}`: The last day q^{end} of the period $q \in \mathcal{Q}_e$ for employee $e \in \mathcal{E}$.
- `q^{b}`: The number of days q^b we prefer to assign employee $e \in \mathcal{E}$ off before the period $q \in \mathcal{Q}_e$ if we assign work in the period.
- `q^{a}`: The number of days q^a we prefer to assign employee $e \in \mathcal{E}$ off after the period $q \in \mathcal{Q}_e$ if we assign work in the period.

```
1 (e,q);q^{start};q^{end};q^{b};q^{a}
2 (e00, 0);2018-09-08;2018-09-09;2;2
3 (e03, 1);2018-09-15;2018-09-16;;2
```

Listing 15: `SurroundedDaysOff`.csv format.

2.17 Patterns

This file contains up to 5 columns, as presented with Listing 16. All entries in this file are patterns $y \in \mathcal{Y}$, corresponding to constraints H9 and S6. The columns are as follows:

- `PatternId`: The id of the pattern.
- `l_{y}`: The length l_y of pattern $y \in \mathcal{Y}$.
- `beta^{pat}_{y}`: A boolean β_y^{pat} denoting whether the pattern should be a hard constraint (H9) or a soft constraint (S6).
- `PatternShifSets`: The shift sets $\langle \sigma_1, \dots \sigma_{l_y} \rangle$ defining pattern $y \in \mathcal{Y}$.
- `omega^{pat}_{y}`: The weight ω_y^{pat} in constraint S6.

```
1 PatternId;l_{y};beta^{pat}_{y};PatternShiftSets;omega^{pat}_{y}
2 1;3;1;[Night, Free, DayAll];
3 6;2;0;[A, Night];150.0
```

Listing 16: `Pattern.csv` format.

2.18 Employee_Date_Pattern

This file contains one column, as presented with Listing 17. All entries are tuples (e, d, y) for $e \in \mathcal{E}, d \in \mathcal{D}^{all}, y \in \mathcal{Y}$ with $\Pi_{e,d,y} = 1$.

```
1 Pi_(e,d,y)=1
2 (e00, '2018-09-12', 4)
3 (e00, '2018-09-12', 6)
```

Listing 17: `Employee_Date_Pattern.csv` format.

2.19 AssignmentLeadsToOther_Ward

This file contains up to 3 columns, as presented with Listing 18. All entries in this file correspond to constraint H13. The columns are as follows:

- `(d1,d2,sigma1,sigma2)`: A tuple $(d_1, d_2, \sigma_1, \sigma_2)$ for a combination of assignments.
- `L^{ward,both}_{d1,d2,sigma1,sigma2}`: A parameter $L^{ward,both}_{d_1,d_2,\sigma_1,\sigma_2}$ denoting whether we can only assign $\sigma_1 \in \mathcal{Z}$ on day $d_1 \in \mathcal{D}$ in combination with assigning $\sigma_2 \in \mathcal{Z}$ on day $d_2 \in \mathcal{D}$, and vice versa.
- `L^{ward,one}_{d1,d2,sigma1,sigma2}`: A parameter $L^{ward,one}_{d_1,d_2,\sigma_1,\sigma_2}$ denoting whether we can only assign $\sigma_1 \in \mathcal{Z}$ on day $d_1 \in \mathcal{D}$ in combination with assigning $\sigma_2 \in \mathcal{Z}$ on day $d_2 \in \mathcal{D}$, but not vice versa.

```
1 (d1,d2,sigma1,sigma2);L^{ward,both}_{d1,d2,sigma1,sigma2}
2 ('2018-09-10', '2018-09-10', DA, AN);1
```

Listing 18: `AssignmentLeadsToOther_Ward.csv` format.

2.20 AssignmentLeadsToOther_Employee

This file contains up to 3 columns, as presented with Listing 19. All entries in this file correspond to constraint H10. The columns are as follows:

- `(e,d1,d2,sigma1,sigma2)`: A tuple $(e, d_1, d_2, \sigma_1, \sigma_2)$ for a combination of assignments for a given employee.
- `L^{emp,both}_{e,d1,d2,sigma1,sigma2}`: A parameter $L^{ward,both}_{d_1,d_2,\sigma_1,\sigma_2}$ denoting whether we can only assign employee $e \in \mathcal{E}$ to $\sigma_1 \in \mathcal{Z}$ on day $d_1 \in \mathcal{D}$ in combination with assigning him to $\sigma_2 \in \mathcal{Z}$ on day $d_2 \in \mathcal{D}$, and vice versa.
- `L^{emp,one}_{e,d1,d2,sigma1,sigma2}`: A parameter $L^{ward,both}_{d_1,d_2,\sigma_1,\sigma_2}$ denoting whether we can only assign employee $e \in \mathcal{E}$ to $\sigma_1 \in \mathcal{Z}$ on day $d_1 \in \mathcal{D}$ in combination with assigning him to $\sigma_2 \in \mathcal{Z}$ on day $d_2 \in \mathcal{D}$, but not vice versa.

```
1 (e,d1,d2,sigma1,sigma2);L^{emp,both}_{e,d1,d2,sigma1,sigma2};L^{emp,one}_{e,d1,d2,sigma1,sigma2}
2 (e00, '2018-09-15', '2018-09-14', AN, N);;1
3 (e00, '2018-09-10', '2018-09-10', ShortDay, Night);;1
```

Listing 19: `AssignmentLeadsToOther_Employee.csv` format.

2.21 Requests

This file contains 3 columns, as presented with Listing 20. We do not use the data in this file when running the optimization, as the information required has already been accounted for in the assignment penalties $\omega^{assign}_{e,d,s}$ as discussed in Section 3.2. Nonetheless, we use the requests to evaluate the quality of the solution in the article, and therefore include them for completeness. The columns are as follows:

- `(e,d)`: A tuple denoting employee $e \in \mathcal{E}$ and day $d \in \mathcal{D}$.
- `Requests`: A set containing tuples with requests employee $e \in \mathcal{E}$ has made for shift assignments day $d \in \mathcal{D}$. The first item in each tuple is a shift set $\sigma \in \mathcal{Z}$ and the second is the priority of the request (i.e., either `High` or `Low`).
- `Non-requests`: This columns is as the previous, except for the shift sets employee $e \in \mathcal{E}$ has requested **not** to be assigned on day $d \in \mathcal{D}$.

```
1 (e,d);Requests;Non-requests
2 (e00, '2018-09-12');{(D, 'Low')};
3 (e00, '2018-09-14');{(Free, 'High')};
4 (e08, '2018-09-25');{(D, 'Low'), (N, 'Low')};
5 (e25, '2018-10-04');;{(A, 'Low')}
```

Listing 20: `Requests.csv` format.

3 Additional information

This section presents some additional information to ensure the reproducibility of the research.

3.1 Shift type ids

Each instance includes various shift types. Different work shifts are denoted with `Dxx`, `Axx` or `Nxx`, depending on whether they occur during day, evening or night shifts, respectively, where `xx` represents numbers. In addition, we have some special shifts, which are as follows:

- `MENT`: The shift type $s_c \in \mathcal{S}^{work} \subset \mathcal{S}$ for chaperoning.
- `PF`: The shift type $s_{pf} \in \mathcal{S}$ for protected days off.
- `AD`: The dummy *free* shift type.
- `BD`: A shift type for parental leave. The shift type is not a work shift, but counts toward the contractual hours.
- `FO`: A shift type for compensatory time off due to public holidays. The shift type is not a work shift, but counts toward the contractual hours.
- `FE`: A shift type for vacation. The shift type is not a work shift, but counts toward the contractual hours.
- `ADM`: A shift type for administrative work.
- `UD` and `UDD`: Shift types for educational activities.
- `NO`: A shift type for days off that count towards the contractual hours. If we require multiple shift types we enumerate them as `NO1`, `NO2`, etc.

3.2 Constraint S1

The weights $\omega^{assign}_{e,d,s}$ in S1 are a combination of various factors, which are as follows:

- A weight associated with assigning a shift type $s \in \mathcal{S}$, and some employees, e.g.:
  - A reward for assigning `PF` during weekends.
  - A reward for assigning trainees to `MENT`.
  - A reward for assigning preferences or a penalty for assigning non-preferences.
  - A reward for assigning specific requests or a penalty for assigning specific requests.
- A weight associated with assigning `FO` $\in \mathcal{S}$ on day $d \in \mathcal{D}$, based on the distance from the corresponding holiday.
- An additional reward when a request is made to certain unpopular shifts, e.g. Friday evening shifts.

We note that to ensure fairness between employees, we have scaled the weights related to preferences and specific requests based on the lowest and the highest objective contribution for each employee.

3.3 Additional parameters and sets

From Appendix A in the article, we have excluded the set $\mathcal{O}_{e,q}$ and the parameter ψ_{d_1,d_2,s_1,s_2}, used for surrounding days off, from the data. We instead obtain them in the following way:

- $\mathcal{O}_{e,q} \subset S^{work} \times D^{all}$: The set of assignments for employee $e \in \mathcal{E}$ that overlap with period $q \in \mathcal{Q}_e$. This set includes all combination of feasible work shifts and days such that assigning $(e, d, s) \in \mathcal{A}$ overlaps in time with the days defined by $\left[q^{start}, q^{end}\right]$.
- ψ_{d_1,d_2,s_1,s_2}: The number of full days between shift assignment $s_1 \in \mathcal{S}^{work}$ on day $d_1 \in \mathcal{D}^{all}$ and $s_2 \in \mathcal{S}^{work}$ on day $d_2 \in \mathcal{D}^{all}$, i.e., $\lfloor T^{start}_{d_2,s_2} - T^{end}_{d_1,s_1} \rfloor$, rounding down in the number of days.

Additionally, due to series constraints, we might need $T^{hours}_{e,d,s}$ for $d \in \mathcal{D}^{pre}$. We only require these parameters for either $s \in \mathcal{S}^{work}$ or $s = \texttt{BD}$. We do not provide these values directly in the data, as we can obtain them in the following way:

- If $s \in \mathcal{S}^{work}$, then $T^{hours}_{e,d,s}$ is the duration of s as given in `ShiftTypes.csv`.
- If $s = \texttt{BD}$, then $T^{hours}_{e,d,s}$ only depends on the employee $e \in \mathcal{E}$. Therefore the value of $T^{hours}_{e,d,s}$ for any day $d \in \mathcal{D}$, given in `FeasibleAssignments.csv`, is valid.