Selecting User Queries in Interactive Job Scheduling[1]

Johannes Varga[a], Günther R. Raidl[a], Tobias Rodemann[b]

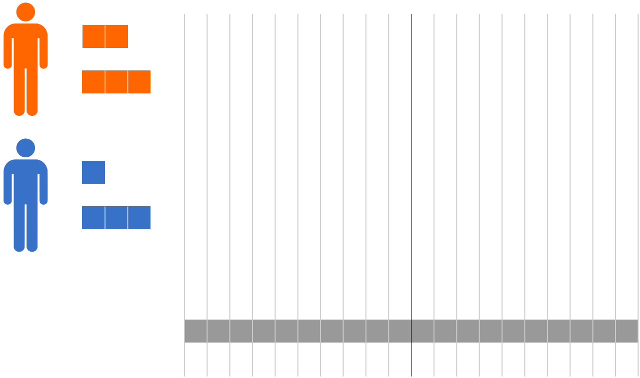
[a]Institute of Logic and Computation, TU Wien, Vienna, Austria

[b]Honda Research Institute Europe, Offenbach, Germany

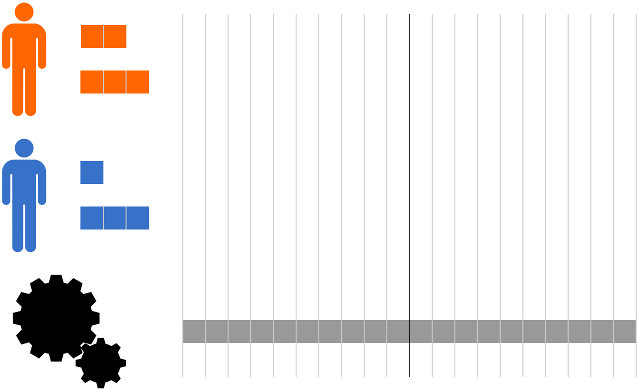
Feb 27, 2024

[1] J. Varga acknowledges the financial support from Honda Research Institute Europe.

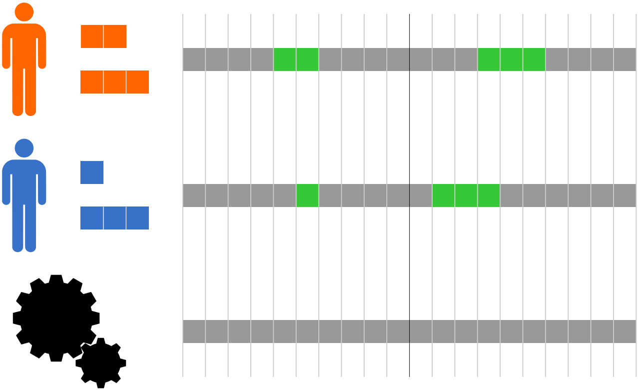
Problem Setting



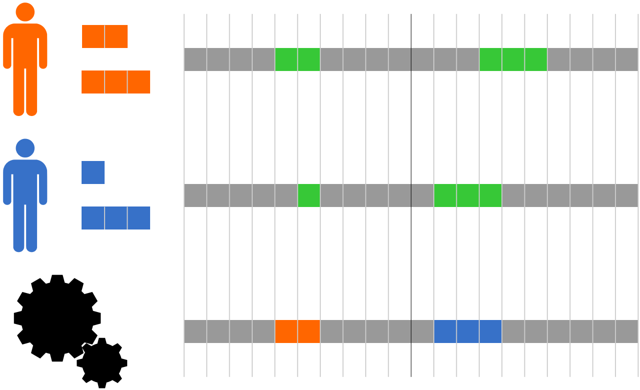
Problem Setting



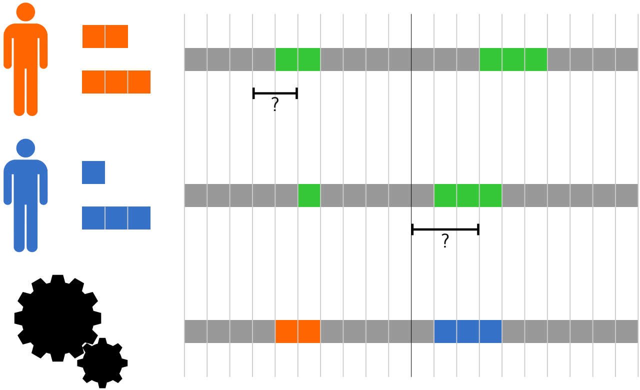
Problem Setting



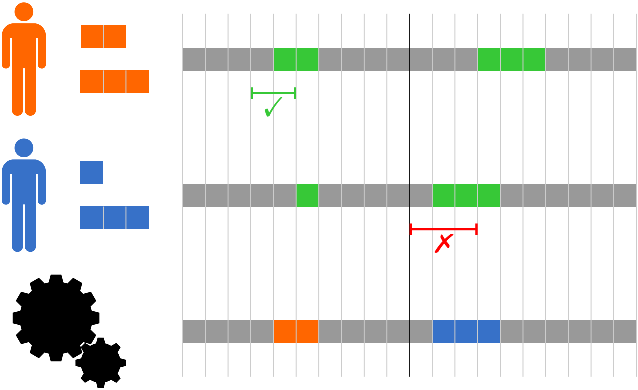
Problem Setting



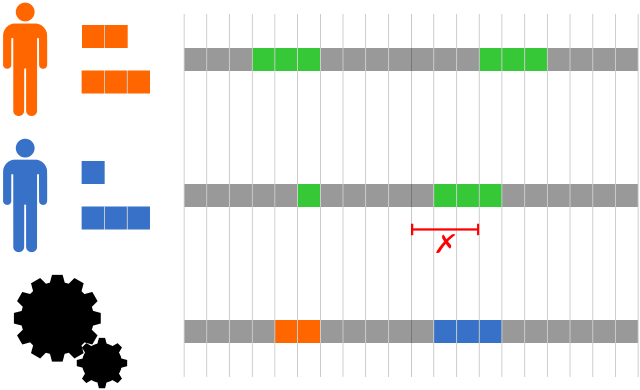
Problem Setting



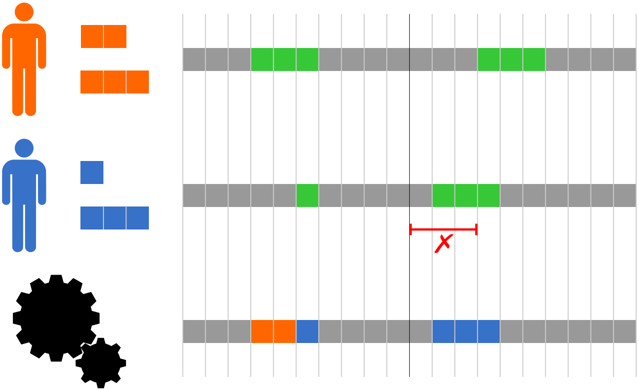
Problem Setting



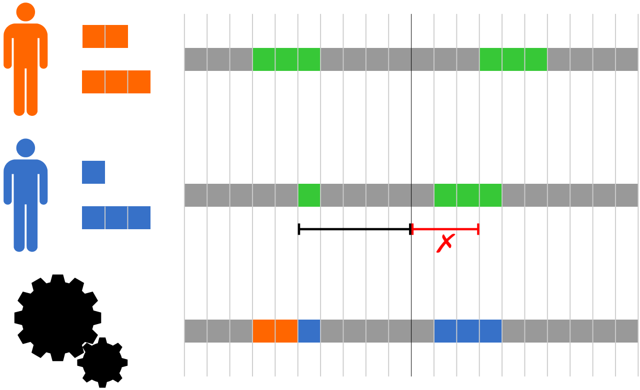
Problem Setting



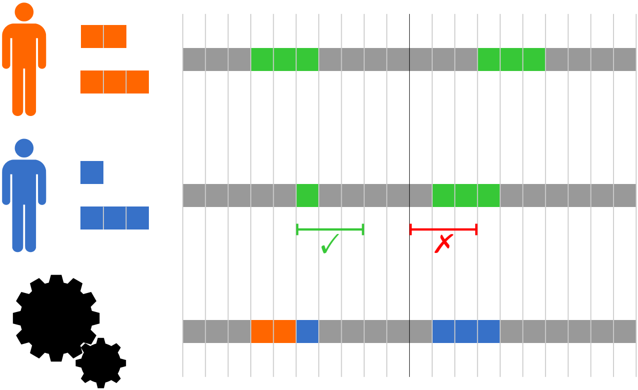
Problem Setting



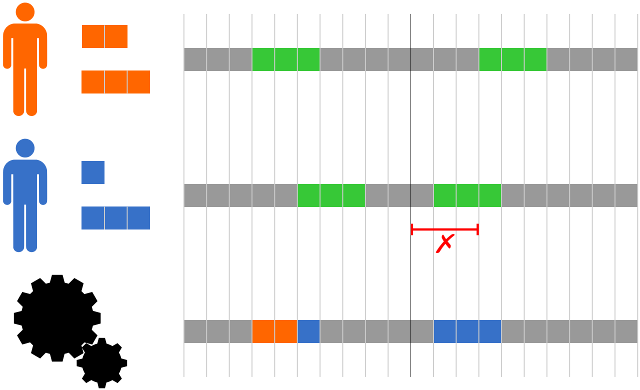
Problem Setting



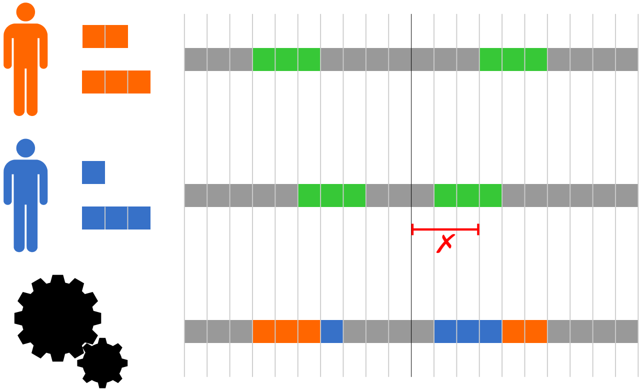
Problem Setting



Problem Setting



Problem Setting



Problem

Core problem:

- Discrete time planning horizon of multiple days
- Multiple users
- Multiple jobs per user
- one machine
- Schedule jobs non-preemptively on machine

Objective:

- time dependent costs c_t for using the machine at timestep t
- penalty q_j for not scheduling a job j
- not scheduling a job is more expensive than scheduling it

Problem

User availabilities:

- Limit job running times
- Only partially known
- Complement knowledge with interaction

Interaction:

- B rounds of interaction
- each with up to b queries
- Query types: Yes/No and Timeframe

Application examples:

- Human resource planning (e.g. scheduling of lectures)
- EV Charging

Literature

Considered problem setting in previous work[2]

Neglect users: Job Scheduling Problem $Pm||TEC$

- MILP and various heuristics (greedy, genetic algorithm, local search) for $Pm||C_{\max}, TEC$[3]
- MILP and Dantzig-Wolfe decomposition for $Rm||TEC$[4]
- Improved MILP[5]

[2] Varga et al. 2023; Varga et al. 2024.
[3] Wang et al. 2018; Anghinolfi, Paolucci, and Ronco 2021.
[4] Ding et al. 2016.
[5] Cheng, Chu, and Zhou 2018; Saberi-Aliabad, Reisi-Nafchi, and Moslehi 2020.

Probabilistic User Models

Criteria for queries:

1. Good response likely $\longrightarrow$ Model users in probabilistic way
2. Improve the schedule $\longrightarrow$ Optimize

Given:

- Known availabilities T^{avail}
- Rejected time intervals I^{rej}

Compute:

- Acceptance probability of queries
- Reasonable user availabilities

Single user, single day

Markov-Model[6]

Model availabilites with Markov process

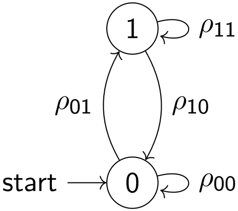


Figure: Two-state Markov Chain

[6] Varga et al. 2023.

Advanced model[7]

Assumptions:

- User available in ≤ 2 intervals
- Morning interval: start 9am$\pm$1h, duration 4h$\pm$1h, probability 90%
- Afternoon interval: start 1pm$\pm$1h, duration 5h$\pm$1h, probability 90%

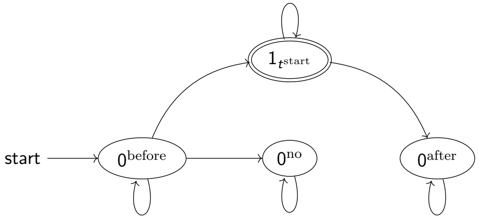
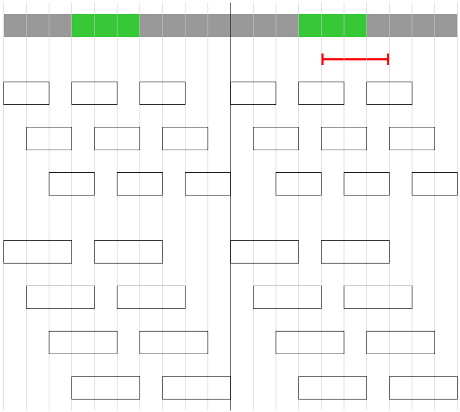


Figure: State diagram for single interval.

[7]Varga et al. 2024.

Probability Threshold[8]



Query candidates

Figure: Procedure in each round

[8]Varga et al. 2023.

Probability Threshold[8]

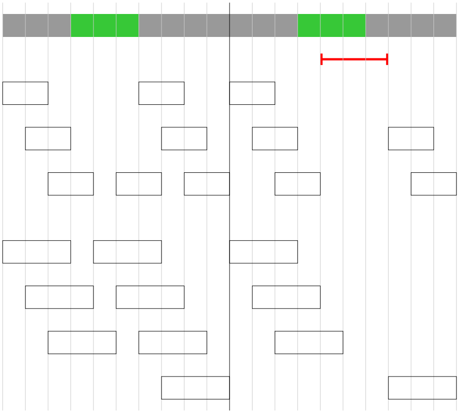
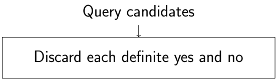
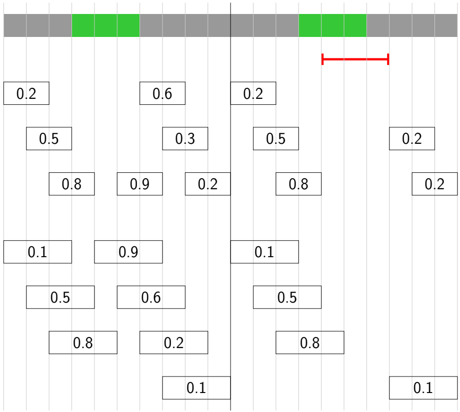


Figure: Procedure in each round



[8] Varga et al. 2023.

Probability Threshold[8]



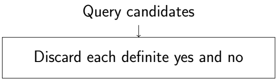


Figure: Procedure in each round

[8]Varga et al. 2023.

Probability Threshold[8]

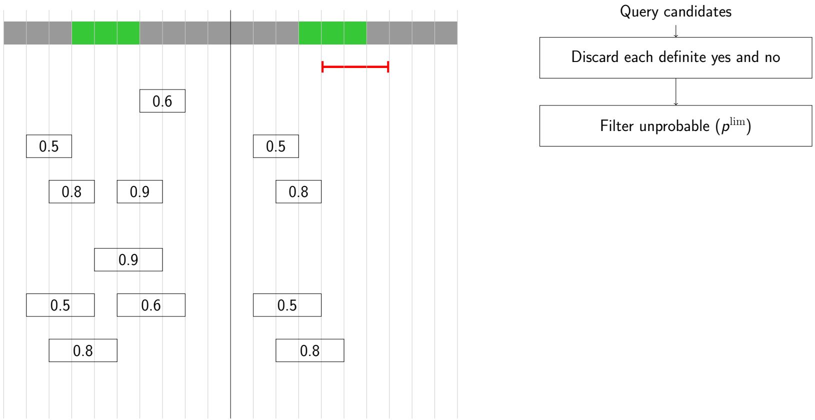
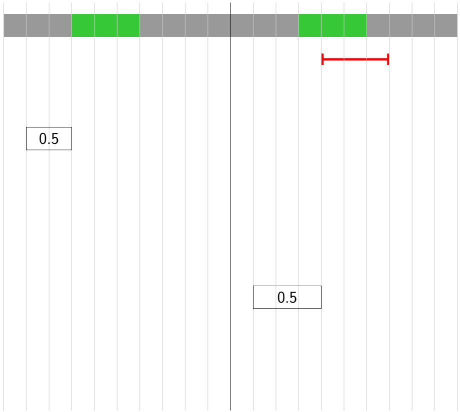


Figure: Procedure in each round

[8]Varga et al. 2023.

Probability Threshold[8]



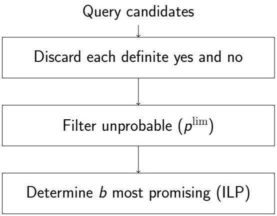
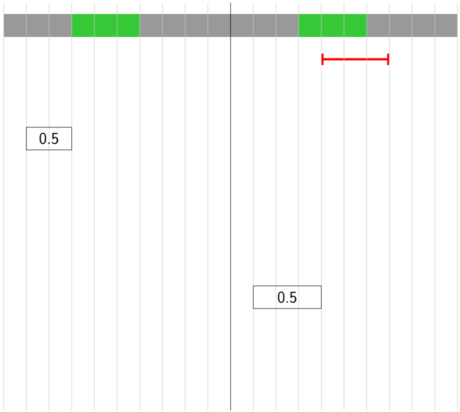


Figure: Procedure in each round

[8] Varga et al. 2023.

Probability Threshold[8]



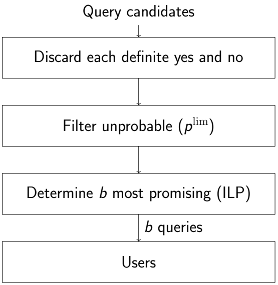


Figure: Procedure in each round

[8] Varga et al. 2023.

Probability Threshold[8]

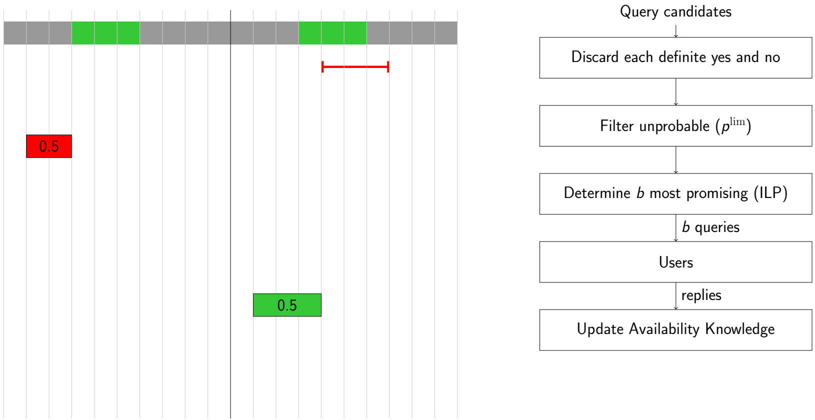


Figure: Procedure in each round

[8] Varga et al. 2023.

Probability Threshold[8]

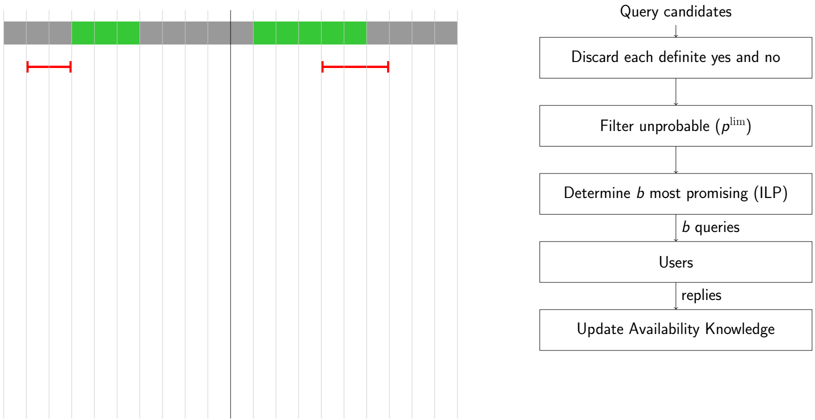
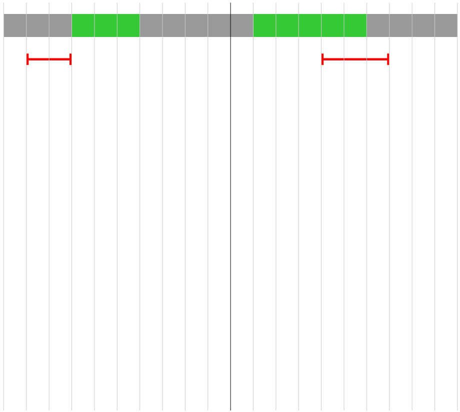


Figure: Procedure in each round

[8]Varga et al. 2023.

Probability Threshold[8]



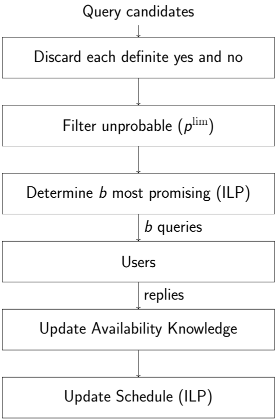


Figure: Procedure in each round

[8]Varga et al. 2023.

Stochastic Programming

Possible queries for each job:

- New starting time
- Half-day timeframe (6am-2pm, 2pm-10pm)

Two stages:

1. Select queries
2. Expected costs

$$
\begin{aligned}
\text{min} \quad & \mathbb{E}_{T^{\text{avail}*}}(\text{ILP}(\mathcal{T}(T^{\text{avail}*}, s^{\text{time}}, s^{\text{frame}}))) \\
\text{s.t.} \quad & \sum_{j \in J} \left(\sum_{t \in T} s_{jt}^{\text{time}} + \sum_{f \in F} s_{jf}^{\text{frame}} \right) \leq b \\
& s_{jt}^{\text{time}} \in \{0,1\} && j \in J, t \in T \\
& s_{jf}^{\text{frame}} \in \{0,1\} && j \in J, f \in F
\end{aligned}
$$

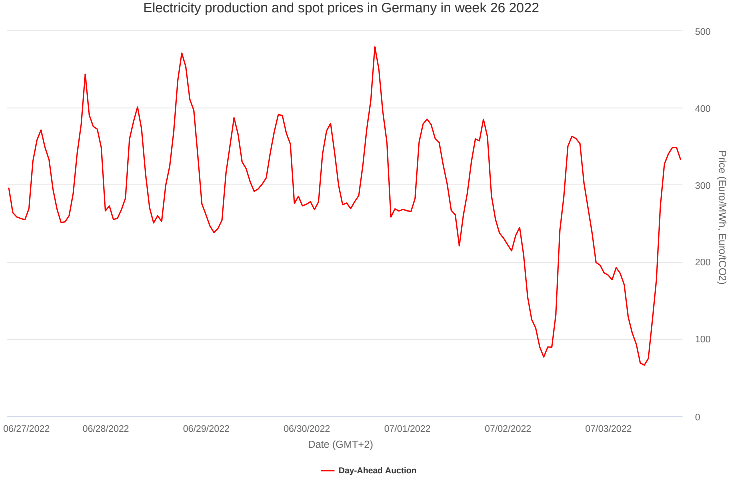
Solve: ILP for sample average approximation (70 samples)

Instance Generation

- one machine, 24 jobs, 6 users
- 30 instances
- five rounds, each with 6 user queries
- five days (6am to 10pm), 4 timesteps per hour
- Job duration: uniformly random from [0.5h,4h]
- Users are simulated and available
  - from 9am$\pm$1h for 4h$\pm$1h with probability 0.9 and
  - from 1pm$\pm$1h for 5h$\pm$1h with probability 0.9
- Known availabilities: one random starting time chosen for each job

Cost Function: Energy Prices in Germany

Job penalty: 40 Euro per timestep



Energy-Charts.info; Data Source: 50 Hertz, Amprion, Tennet, TransnetBW, EEX, EPEX SPOT; Last Update: 01/04/2023, 2:08 PM GMT+1

Results 1

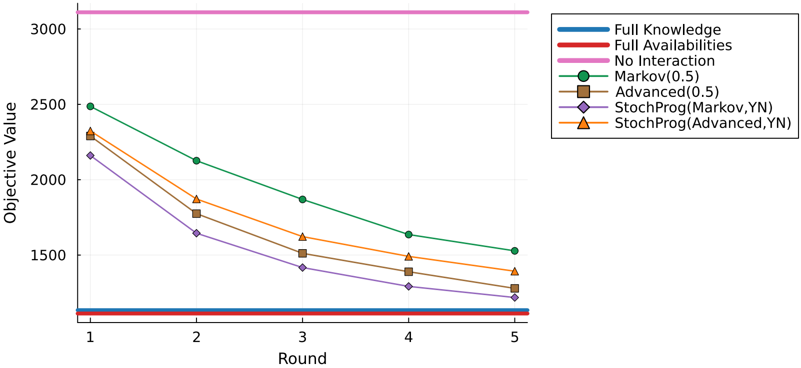


Figure: Mean objective values, 24 jobs, 6 users

Results 2

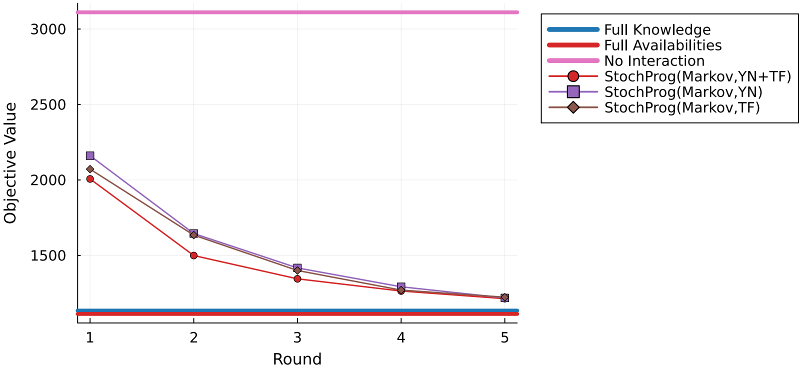


Figure: Mean objective values, 24 jobs, 6 users

Results 3

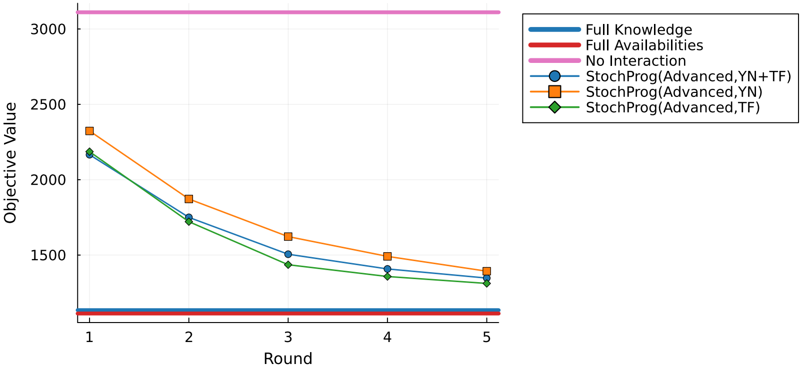


Figure: Mean objective values, 24 jobs, 6 users

Conclusion and Future Work

Conclusions:

- Stochastic programming significantly better than probability threshold approach
- But: harder to solve
- Query type: small influence
- Markov model better for stochastic programming

Future Work:

- Learn user model from historic availability data
- Heuristic and/or Benders decomposition → solve larger instances

Thank you!

ILP (Interactive Job Scheduling)

$$
\begin{aligned}
\mathrm{ILP}(\mathcal{T}) \quad \min \quad & \sum_{j\in J}\sum_{t\in\mathcal{T}_j}\sum_{t'\in T_j[t]} c_{t'} x_{jt} + \sum_{j\in J} q_j \left(1 - \sum_{t\in\mathcal{T}_j} x_{jt}\right) \\
\text{s.t.} \quad & \sum_{t\in\mathcal{T}_j} x_{jt} \le 1 && j \in J \\
& \sum_{j\in J}\ \sum_{t\in\mathcal{T}_j \mid t'\in T_j[t]} x_{jt} \le 1 && t' \in T \\
& \sum_{j\in J_u}\ \sum_{t\in\mathcal{T}_j \mid t'\in T_j[t]} x_{jt} \le 1 && u \in U,\ t' \in T \\
& x_{jt} \in \{0,1\} && j \in J,\ t \in \mathcal{T}_j
\end{aligned}
$$

J ... Jobs $\quad$ T ... Time horizon

$\mathcal{T}_j$ ... allowed starting times for job j

x_{jt} ... 1, iff job j starts at timestep t

$T_j[t]$ ... timesteps job j runs in when started at timestep t

Second Stage of Probabilistic Program

n^{samples} instances of ILP(T^{job}) with variables $x_{jt}^{(k)}$ and additional constraints

$$x_{jt}^{(k)} = 0 \qquad j \in J, t \in T \setminus T_j^{\mathrm{pos,(k)}}$$

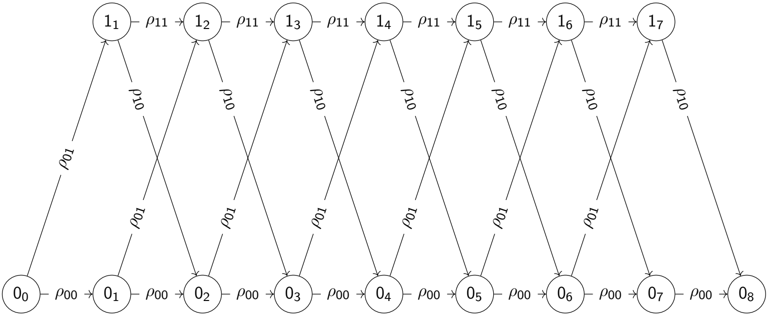
$$s_{jt}^{\mathrm{time}} + \sum_{f \in F_j^{\mathrm{pos,(k)}} \mid t_{jf}^{\mathrm{reply,(k)}} = t} s_{jf}^{\mathrm{frame}} \geq x_{jt}^{(k)} \qquad u \in U, j \in J_u, t \in T, T_j[t] \not\subseteq T_u^{\mathrm{avail}}$$

Computing Environment

- Julia 1.9.3
- Gurobi 10.0.3 (single-threaded) via JuMP
- Single core of Intel Xeon E5-2640 v4
- 60min timelimit for Gurobi

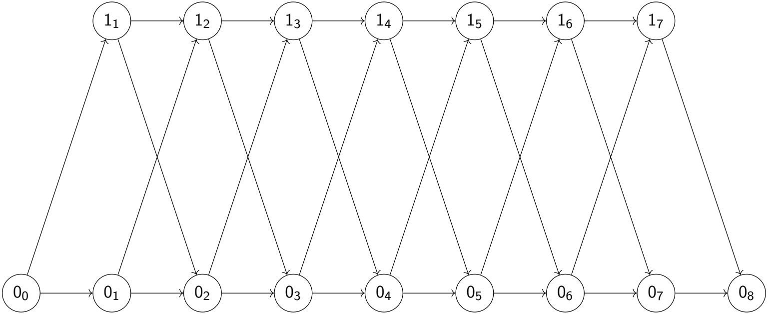
Probability Calculation: Graph

E.g. day with 7 timesteps, $T^{\text{avail}} = \{5, 6\}$, $I^{\text{rej}} = \{[1, 5], [3, 7]\}$



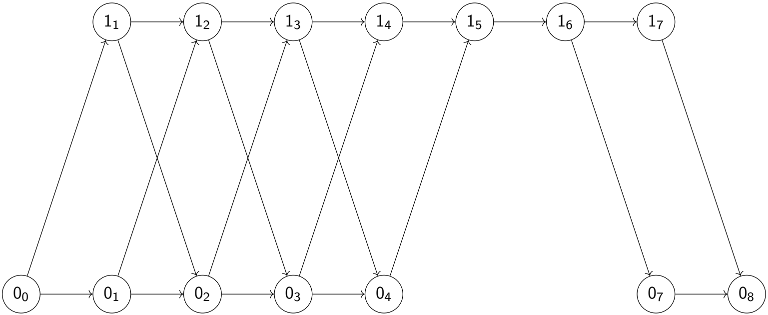
Probability Calculation: Graph

E.g. day with 7 timesteps, $T^{\text{avail}} = \{5, 6\}$, $I^{\text{rej}} = \{[1, 5], [3, 7]\}$



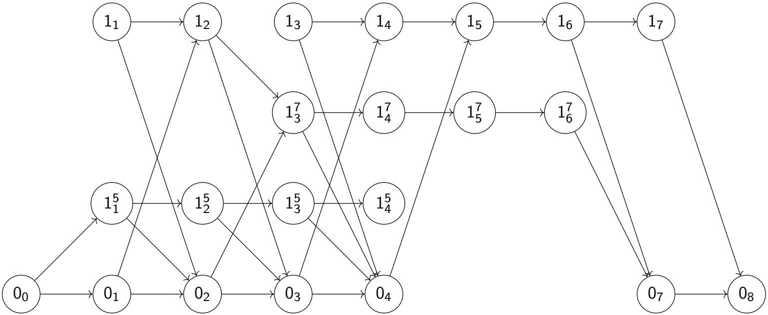
Probability Calculation: Graph

E.g. day with 7 timesteps, $T^{\text{avail}} = \{5, 6\}$, $I^{\text{rej}} = \{[1, 5], [3, 7]\}$



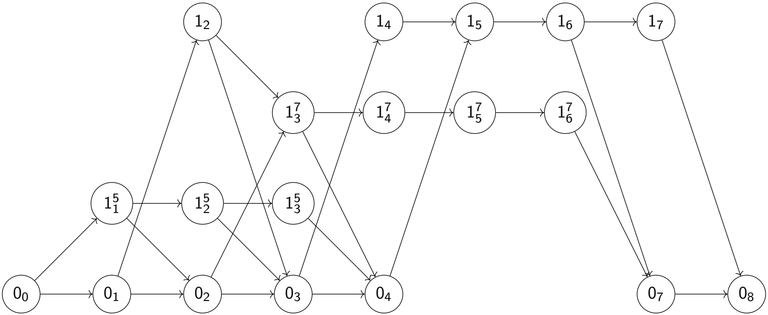
Probability Calculation: Graph

E.g. day with 7 timesteps, $T^{\text{avail}} = \{5, 6\}$, $I^{\text{rej}} = \{[1, 5], [3, 7]\}$



Probability Calculation: Graph

E.g. day with 7 timesteps, $T^{\text{avail}} = \{5, 6\}$, $I^{\text{rej}} = \{[1, 5], [3, 7]\}$



Probability Calculation

$$p^{\text{path}}_{0_0,v} = \sum_{P \in \text{Paths}(0_0,v)} \prod_{(w,w') \in P} \rho(w,w')$$

$$= \sum_{u \in N^-(v)} \sum_{P \in \text{Paths}(0_0,u)} \left(\prod_{(w,w') \in P} \rho(w,w') \right) \cdot \rho(u,v)$$

$$= \sum_{u \in N^-(v)} p^{\text{path}}_{0_0,u} \rho(u,v)$$

$$p^{\text{path}}_{v,0_{t^{\max}+1}} = \sum_{w \in N^+(v)} p^{\text{path}}_{w,0_{t^{\max}+1}} \rho(v,w)$$

Probability Calculation

Probability that $[\tau, \tau']$ will be accepted:

$$p^{\mathrm{avail}}([\tau, \tau'] \mid T^{\mathrm{avail}}, I^{\mathrm{rej}}, 0_{t^{\max}+1}) = \frac{\sum_{P \in \text{1-Paths}(\tau,\tau')} p^{\mathrm{path}}_{0_0, P_\tau} \cdot \rho_{11}^{\tau'-\tau} \cdot p^{\mathrm{path}}_{P_{\tau'}, 0_{t^{\max}+1}}}{p^{\mathrm{path}}_{0_0, 0_{t^{\max}+1}}}$$

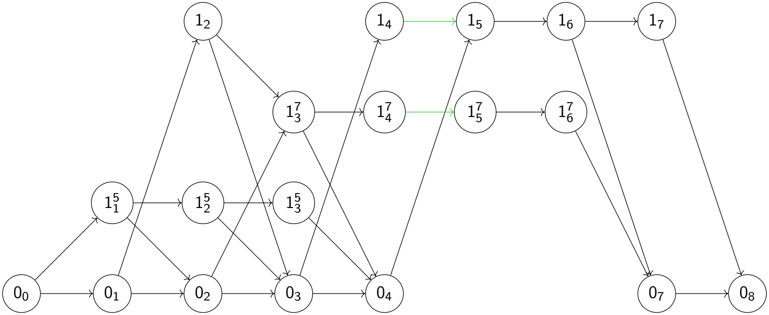


Figure: 1-Paths(4,5) for the example in green

Results 4

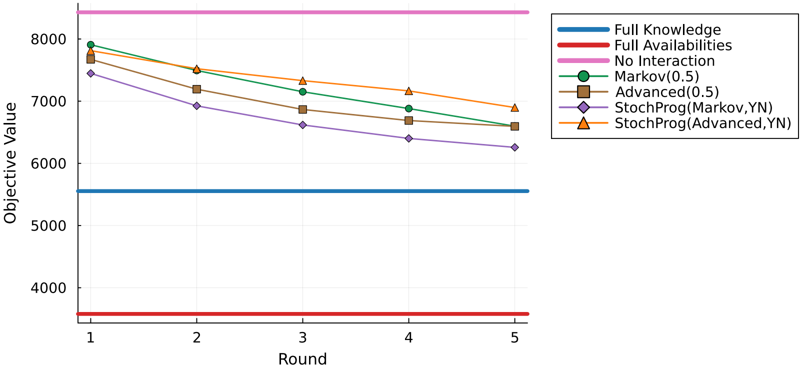


Figure: Mean objective values, 48 jobs, 12 users

Results 5

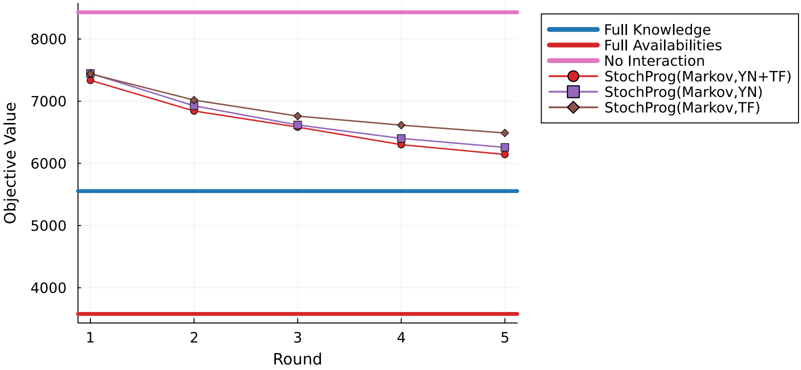


Figure: Mean objective values, 48 jobs, 12 users

Results 6

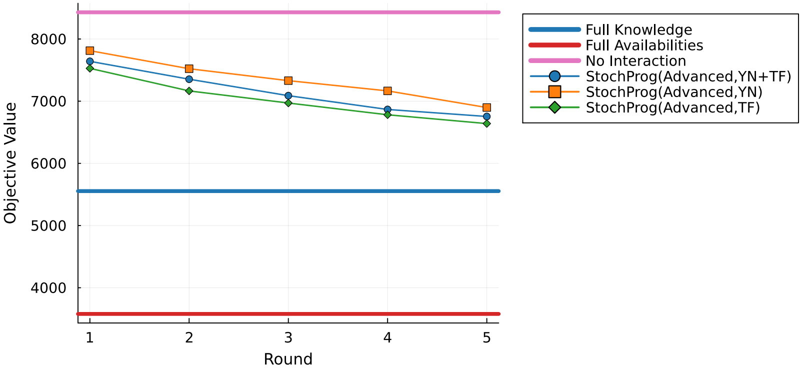


Figure: Mean objective values, 48 jobs, 12 users