An interactive optimization method to promote ethics for Nurse Rostering

Vincent Bebien[1,2], Odile Bellenguez[1], Gilles Coppin[3], Anna Ma-Wyatt[2], and Rachel Stephens[2]

[1] IMT Atlantique, La Chantrerie, 4 rue Alfred Kastler, 44307, Nantes, France
[2] University of Adelaide, School of Psychology, SA 5005, Adelaide, Australia
[3] IMT Atlantique, Lab-STICC, Technopôle Brest-Iroise CS 83818, 29238, Brest, France

Keywords: Nurse Rostering Problem, Interactive methods, AI Ethics.

1 Introduction

In the field of AI Ethics, scholars have identified various kinds of ethical issues related to autonomous decision-making algorithms [6], which include Operations Research (OR) applications. While methods for addressing some ethical issues have been studied in scheduling contexts, especially fairness criteria [10], there are still research avenues for OR applications that have received limited attention. We consider that ethics cannot be efficiently integrated into a decision tool without considering the specific and dynamic aspects of the problem on the field [2]; thus we propose here a design for a decision tool for the Nurse Rostering Problem (NRP) that allows for a better integration of moral values and ethical considerations.

2 Computing moral values

In the field of ethics, multiple frameworks have been developed to describe the moral preferences of individuals by identifying core values one would wish to respect. Taking into account one of them, the basic human values theory [8], we try to address the main issues of NRP by identifying first which aspects may be related to which of the theory's ethical principles. For example, some constraints such as satisfying minimum personnel requirements may be related to the *conformity* value, while others such as balancing workload across employees can be considered as *benevolence* and *universalism*. This approach forms a basis for a moral compass of decision-makers.

Mathematically, these potentially conflicting values are represented with norms that may be either modeled as objective functions or constraints. Hard constraints may be used to represent a threshold that has to be attained regarding a certain norm, refusing all solutions that do not meet it. Alternatively, using soft constraints allows the consideration of such solutions as valid but of lesser fitness, depending on their assigned weights.

These weights implicitly create a hierarchy between soft constraints, where the ones with the highest penalties will be preferred to the others. Thus, an 'ethical profile' can be drawn from the way the objective function is modeled. Multi-criteria decision-making (MCDM) methods allow users to visualize the different set objectives and/or trade-offs

between efficient solutions. For these methods, all objectives are considered equivalent in the model; only the end-user has the agency to either choose their preferred solution or decide how these objectives should be ranked or prioritized.

An important limitation of standard modeling approaches with hard and soft constraints with a single objective function when it comes to ethical decision-making is its static aspect. As moral values are constantly evolving [9], a mathematical model prioritizing some criteria and enforcing ethical constraints might become irrelevant and unreliable in the future, or for people with a different cultural background [1]. MCDM approaches such as interactive methods [4] offer some flexibility by including the decision-maker in the loop, allowing them to decide which criteria are most important in their current situation. Nevertheless, these criteria themselves as well as the problem structure typically remain the same, cannot be modified and might also lose relevance with time passing and context changing.

We argue that a human-in-the-loop decision-making process that gives more agency than standard MCDM interactive methods could be used to build a tool that better considers ethics. An interactive process offers some advantages that could benefit the whole nurse scheduling process. Incorporating human interaction allows for a better adaptation to new conditions, which helps to generate well-suited schedules and reinforces user agency as they may have a better comprehension of the whole process [7]. An open process also allows other stakeholders such as nurses to better grasp how a schedule has been designed, which might be regarded as a fair process [3].

3 Integrating ethical considerations into an interactive tool

We propose here to use an interactive reoptimization method [5] adapted to an NRP, where the decision-maker can iteratively modify the set of rules, which correspond to hard constraints, to obtain new solutions. These modifications may consist of either local changes (e.g. assigning a nurse to a certain shift on a specific day) or global changes (e.g. forbidding some shift patterns for all nurses). More specifically, users have access to a catalogue of 'template' rules that can be parameterized according to their preferences. For example, the catalogue contains a template called 'Limit consecutive working days' that can be parameterized by selecting the nurses and period for which this rule should be applied, as well as the limit value. Whenever a rule is added to (or deleted from) the model, a new solution is generated according to the changes, following a user-defined optimization criterion also chosen from a catalogue.

This design aims to provide an interactive tool that displays and allows changes to the main aspects of the mathematical models that are used. While the preliminary work that established the catalogue limits the decision-maker's possibilities, it allows non-experts in OR to directly manipulate the NRP formulation. The proposed design may be especially useful when a clash between two ethical criteria arises and arbitration is needed to obtain a feasible solution. Through trial and error, the user may have a better understanding of the problem structure and the different trade-offs they should consider. While we focus on the scheduling process itself, such tool could also be used in a reoptimization context, when unplanned events may arise during the scheduled period.

A possible drawback of this method is that ethical aspects related to scheduling might be ignored or forgotten in the process, as scheduling tasks are often difficult for human decision makers. To help the user detect potential ethical flaws in a candidate schedule, we propose to implement the presentation of specific ethical recommendations that would highlight some of them. The set of presented recommendations may be determined by the user's ethical preferences, which could be assessed either beforehand or during the iterative process. This information could be used either to show recommendations that are preferred by the user or to nudge them towards other ethical criteria they would otherwise not likely consider.

References

1. Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.F., Rahwan, I.: The moral machine experiment. Nature **563**(7729), 59–64 (2018). https://doi.org/10.1038/s41586-018-0637-6
2. Bebien, V., Bellenguez, O., Coppin, G., Ma-Wyatt, A., Stephens, R.: Ethical decision-making in human-automation collaboration: a case study of the nurse rostering problem. AI and Ethics (In press), https://hal.science/hal-04500402
3. Greenberg, J.: Organizational justice: Yesterday, today, and tomorrow. Journal of Management **16**(2), 399–432 (1990). https://doi.org/10.1177/014920639001600208
4. Korhonen, P.: Interactive methods. Multiple criteria decision analysis: state of the art surveys pp. 641–661 (2005). https://doi.org/10.1007/0-387-23081-5_16
5. Meignan, D., Knust, S., Frayret, J.M., Pesant, G., Gaud, N.: A review and taxonomy of interactive optimization methods in operations research. ACM Transactions on Interactive Intelligent Systems (TiiS) **5**(3), 1–43 (2015). https://doi.org/10.1145/2808234
6. Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S., Floridi, L.: The ethics of algorithms: Mapping the debate. Big Data & Society **3**(2), 2053951716679679 (2016). https://doi.org/10.1177/2053951716679679
7. Onnasch, L., Wickens, C.D., Li, H., Manzey, D.: Human performance consequences of stages and levels of automation: An integrated meta-analysis. Human Factors **56**(3), 476–488 (2014). https://doi.org/10.1177/0018720813501549
8. Schwartz, S.H.: Are there universal aspects in the structure and contents of human values? Journal of Social Issues **50**(4), 19–45 (1994). https://doi.org/10.1111/j.1540-4560.1994.tb01196.x
9. Smith, K.B., Alford, J.R., Hibbing, J.R., Martin, N.G., Hatemi, P.K.: Intuitive ethics and political orientations: Testing moral foundations as a theory of political ideology. American Journal of Political Science **61**(2), 424–437 (2017). https://doi.org/10.1111/ajps.12255
10. Wolbeck, L.A.: Fairness aspects in personnel scheduling. Discussion Papers 2019/16, Free University Berlin, School of Business & Economics (2019). https://doi.org/10.17169/refubium-26050