European Journal of Operational Research 239 (2014) 89–101

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Discrete Optimization

Multiobjective shortest path problems with lexicographic goal-based preferences

Francisco Javier Pulido [1], Lawrence Mandow *, José Luis Pérez de la Cruz

Dpto. Lenguajes y Ciencias de la Computación, Universidad de Málaga, Boulevar Louis Pasteur, 35. Campus de Teatinos, 29071 Málaga, Spain

ARTICLE INFO

Article history:
Received 26 April 2013
Accepted 4 May 2014
Available online 15 May 2014

Keywords:
Combinatorial optimization
Multiobjective shortest path problem
Label-setting search
Heuristic search
Goal programming

ABSTRACT

Multiobjective shortest path problems are computationally harder than single objective ones. In particular, execution time is an important limiting factor in exact multiobjective search algorithms. This paper explores the possibility of improving search performance in those cases where the interesting portion of the Pareto front can be initially bounded. We introduce a new exact label-setting algorithm that returns the subset of Pareto optimal paths that satisfy a set of lexicographic goals, or the subset that minimizes deviation from goals if these cannot be fully satisfied. Formal proofs on the correctness of the algorithm are provided. We also show that the algorithm always explores a subset of the labels explored by a full Pareto search. The algorithm is evaluated over a set of problems with three objectives, showing a performance improvement of up to several orders of magnitude as goals become more restrictive.

1. Introduction

Goal programming is one of the most successful models of Multicriteria Decision Theory (Chankong & Haimes, 1983). Virtually hundreds of applications can be found in the literature (Romero, 1991; Tamiz, Jones, & El-Darzi, 1995). This paper explores the application of the goal-based decision paradigm to multicriteria shortest path problems.

Multicriteria shortest path problems arise naturally in many fields, such as robot surveillance (Delle Fave, Canu, Iocchi, Nardi, & Ziparo, 2009), robot path planning (Fujimura, 1996), satellite scheduling (Gabrel & Vanderpooten, 2002), and route planning in different contexts (Delling & Wagner, 2009; Clímaco, Craveirinha, & Pascoal, 2003; Jozefowiez, Semet, & Talbi, 2008; Machuca & Mandow, 2012). A number of shortest path algorithms have been proposed to tackle different multicriteria decision models. The work of Hansen (Hansen, 1979) presented a bi-objective extension of Dijkstra's label setting algorithm. Martins (Martins, 1984) proposed a general multiobjective label setting algorithm. A recent evaluation of several multiobjective shortest path algorithms can be found in (Raith & Ehrgott, 2009).

The multiobjective shortest path problem is computationally harder than the single objective one. The number of label expansions can grow exponentially with solution depth, even for the two objective case (Hansen, 1979). With the assumption of bounded integer costs and a fixed number of objectives the problem becomes tractable for polynomially sized graphs, but still harder than single objective search (e.g. see (Mandow & Pérez de la Cruz, 2009; Müller-Hannemann & Weihe, 2006)).

Search efficiency can be improved in single destination (one to one) problems using lower bound distance estimates in a similar way as algorithm A* improves over Dijkstra's (Pearl, 1984). Several multiobjective extensions of A* have been proposed. These can be grouped in two classes: those that perform *node expansion* as its basic operation (like MOA* (Stewart & White, 1991)), and those that perform *label expansion* (like Tung and Chew's algorithm (Tung & Chew, 1992) and NAMOA* (Mandow & Pérez de la Cruz, 2010)). The interest in these algorithms with lower bounds is justified by the fact that: precise lower bound estimates can be efficiently precalculated for a large class of problems (Tung & Chew, 1992); and the use of such estimates still guarantees an exact solution, i.e. the algorithms find the set of all Pareto optimal solutions to the problem.

Several algorithms extended the node expansion policy of MOA* to different contexts, like algorithms MOA** for search with non-consistent lower bounds (Dasgupta, Chakrabarti, & DeSarkar, 1999), BCA* for compromise solutions (Galand & Perny, 2006), or METAL-A* for goal based preferences (Mandow & Pérez de la Cruz, 2001). The latter are the subject of this work. However, recent empirical and formal analyses (Machuca, Mandow, Pérez de la Cruz, & Ruiz-Sepulveda, 2012; Pérez de la Cruz, Mandow, &

* Corresponding author.
E-mail addresses: francis@lcc.uma.es (F.J. Pulido), lawrence@lcc.uma.es (L. Mandow), perez@lcc.uma.es (J.L. Pérez de la Cruz).
[1] Main corresponding author. Tel.: +34 (9)52 132863.

http://dx.doi.org/10.1016/j.ejor.2014.05.008
0377-2217/

Machuca, 2013) have shown that lower bounded search with node expansion can perform much worse than blind search algorithms and, more precisely, that performance can seriously degrade with better lower bound estimates. In practice, this result ruins the primary purpose of using lower bounds in these algorithms in the first place.

On the other hand, label expansion algorithms with lower bounds have successfully improved performance over blind search algorithms. The efficiency of NAMOA* has been formally shown to improve with better informed lower bound estimates and, in fact, it has been shown to optimally exploit such estimates among the class of admissible algorithms (Mandow & Pérez de la Cruz, 2010). Empirical results confirm that NAMOA* performs consistently better than blind search, and that better informed lower bounds result in faster search with less space requirements (Machuca et al., 2012). Experiments on problems like bicriteria route planning reveal that time, rather than space, is the practical limiting factor in the calculation of the full Pareto set of solutions (Machuca & Mandow, 2012; Machuca, Mandow, & Pérez de la Cruz, 2009). Recent attempts to improve this algorithm include parallel search (Sanders & Mandow, 2013) and the use of specific efficient data structures (Mali, Michail, & Zaroliagis, 2012).

Many problems do not require in practice the calculation of the full Pareto optimal set of solutions. In this work we investigate the possibility of further improvements over the efficiency of NAMOA* through the introduction of lexicographic goal based preferences. A set of goals can be proposed to bound the area of interesting solutions. More precisely, given a set of goals, we tackle the problem of finding the subset of Pareto optimal paths that satisfy the goals or, if these cannot be satisfied, finding the subset of Pareto optimal paths that minimize deviation from the goals. We propose a new multicriteria label-setting algorithm with lower bounds and label expansion that finds such goal-optimal solutions. The new algorithm explores a subset of the labels explored by NAMOA*, achieving important performance improvements.

Section 2 reviews relevant concepts from multicriteria decision theory and introduces the concept of pruning preference. Section 3 describes the algorithm. Important properties concerning admissibility and efficiency are presented in Section 4. An empirical evaluation is described and discussed in Section 5. Finally some conclusions and future work are outlined.

2. Preliminaries

2.1. Lexicographic goal preferences

First of all, we review the concepts of attribute, objective, and goal, as defined in Romero (1991). Let X be the set of solutions to a decision problem. An **attribute** is a measurable property $g(x) : X \to \mathbb{R}$. An **objective** represents the desired improvement of an attribute, i.e. maximization or minimization. A **goal** combines an attribute with a specific *target* value, or aspiration level $t \in \mathbb{R}$, stated by the decision maker to define his/her preference. Goals for multiobjective shortest path problems are always of the form $g(x) \leqslant t$. Goals are not constraints, i.e. feasible solutions may not satisfy all goals.

Let us consider a set of q attributes $g_i : X \to \mathbb{R}, 1 \leqslant i \leqslant q$ grouped in l priority levels sorted in order of decreasing preemptive importance. Each priority level k comprises a set I_k of one or more attributes. Goals are defined by setting *targets* t_i for each attribute, $g_i(x) \leqslant t_i$.

A solution to a goal problem is *satisfactory* when *all* the goals can be satisfied. We seek nondominated satisfactory solutions. If there are no satisfactory solutions to a problem, we seek nondominated solutions that minimize deviation from the targets. In lexicographic goal problems, the deviation of a set of goals is measured separately for each priority level. Minimizing the deviation of goals at level k is infinitely more important than minimizing deviation at level $k+1$.

Several methods have been proposed to measure the deviation from a set of goals. In this work, the minimization of the weighted sum of deviations is employed. Let $\vec{g} = (g_1, g_2, \ldots, g_q)$ be a vector with all attributes (costs) of a given solution $x \in X$. We can calculate a deviation vector for $\vec{g}$ with one component for each priority level, $\vec{d}(\vec{g}) = (d_1(\vec{g}), d_2(\vec{g}), \ldots, d_l(\vec{g}))$. For each level k, its deviation d_k can be defined as:

$$d_k(\vec{g}) = \sum_{i \in I_k} w_i \times \max(0, g_i - t_i) \tag{1}$$

where w_i is the relative weight of goal i in level k.

We define the *optimum achievement* vector $\vec{d}^* = (d_1^*, d_2^*, \ldots, d_l^*)$ as the minimum lexicographic deviation vector among all solutions. Thus, the set of goal-optimal solutions consists of all nondominated feasible solutions with a deviation equal to $\vec{d}^*$. If there is a satisfactory solution, then the optimum achievement vector is equal to $\vec{0}$.

2.2. Formal definitions

We will now reproduce some standard definitions and introduce some new preference relations between cost vectors $\vec{y}, \vec{y}' \in \mathbb{R}^q$.

- *Dominance* ($\prec$) or Pareto-optimal preference is defined as follows,

$$\vec{y} \prec \vec{y}' \iff \forall i \quad y_i \leqslant y'_i \wedge \vec{y} \neq \vec{y}' \tag{2}$$

Dominance is a strict partial order. Given a set of vectors X, we shall define $\mathcal{N}(X)$ the set of nondominated vectors in set X in the following way,

$$\mathcal{N}(X) = \{\vec{x} \in X \mid \nexists \vec{y} \in X \quad \vec{y} \prec \vec{x}\} \tag{3}$$

We shall find it useful to denote by $\preceq$ the relation "dominates or equals".

- Let us denote $\alpha_i = \min_{\vec{x} \in \mathcal{N}(X)}\{x_i\}$, and $\beta_i = \max_{\vec{x} \in \mathcal{N}(X)}\{x_i\}$. The set $\mathcal{N}(X)$ is bounded by the *ideal point* $\vec{\alpha} = (\alpha_1 \ldots \alpha_q)$, and the *nadir point* $\vec{\beta} = (\beta_1 \ldots \beta_q)$. The ideal point can be calculated optimizing each objective separately. However, for $q > 2$ it is difficult to calculate the nadir point without computing the whole set of nondominated solutions.
- *Lexicographic order* $\prec_L$ is defined as follows,

$$\vec{y} \prec_L \vec{y}' \iff \exists j \; y_j < y'_j \wedge \forall i < j \quad y_i = y'_i \tag{4}$$

The lexicographic order is a strict total order. The lexicographic optimum of a set of vectors is trivially a nondominated vector.

- We define *lexicographic goal* preferences ($\prec_G$) as a partial order relation,

$$\vec{y} \prec_G \vec{y}' \iff \vec{d}(\vec{y}) \prec_L \vec{d}(\vec{y}') \vee (\vec{d}(\vec{y}) = \vec{d}(\vec{y}') \wedge \vec{y} \prec \vec{y}') \tag{5}$$

It is easy to see that $\prec_G$ is a strict partial order (it is irreflexive and transitive). Given a set of vectors X, we shall define $\mathcal{O}_G(X)$ the set of optimal vectors in X according to lexicographic goal preferences (i.e. *goal-optimal* vectors) as,

$$\mathcal{O}_G(X) = \{\vec{x} \in X \mid \nexists \vec{y} \in X \quad \vec{y} \prec_G \vec{x}\} \tag{6}$$

Notice that an optimal solution according to $\prec_G$ is also a nondominated solution, i.e. $\mathcal{O}_G(X) \subseteq \mathcal{N}(X)$.

- Let us consider a goal $y_k \leqslant t_k$, the *slack variable* s_k for this goal is defined as