Journal of the Operations Research
Society of Japan
Vol. 37, No. 3, September 1994

AN APPROACH FOR WORST CASE ANALYSIS OF HEURISTICS: ANALYSIS OF A FLEXIBLE 0-1 KNAPSACK PROBLEM

Tsung-Chyan Lai
National Taiwan University

Margaret L. Brandeau Samuel Chiu
Stanford University

(Received February 5, 1993)

Abstract We introduce one technique for worst case analysis of heuristics. We use functions of continuous variables determined by the input data to represent the ratio of the solution value generated by an approximate algorithm(or a lower bound on it) over an upper bound on the optimal solution value. We then use standard mathematical techniques to analyze the function corresponding to the performance ratio. By taking the infimum of such function over all problem instances, the (tight) worst-case performance ratio of the algorithm is thus obtained. To illustrate the approach, we analyze a flexible 0-1 knapsack problem.

1. Introduction

The increased focus on heuristics for solving large-scale NP-hard [3] problems has led to interest in evaluating–and improving–their performance. In addition to serving as one criterion for evaluating the performance of a heuristic, worst case analysis, especially in a data-dependent setting, can provide insights into the behavior of a heuristic [1]. In this paper we use an optimization approach for deriving a worst case performance ratio of a heuristic. This approach may be particularly useful for analysis of compound heuristics, where two or more heuristics are applied separately and the best solution is chosen [2], [4], [7]. When the individual heuristics complement one another, the worst case performance of the compound heuristic can be superior to that of the individual heuristics.

Without loss of generality, we assume our problem to be of the maximization type. We let Z^H represent the objective value obtained by using heuristic H, Z^* be the optimal solution, and η_H be its associated worst case performance ratio, i.e.,

$$\eta_H = \inf_I \{\frac{Z^H(I)}{Z^*(I)}\} \qquad \text{for all problem instances } I,$$

where $Z^H(I)$ and $Z^*(I)$ are the corresponding solutions given the problem instance I.

Since an explicit form for the optimal solution, $Z^*(I)$, is typically not available, we introduce an attainable upper bound, $Z^{S*}(I)$, as a substitute. If an explicit form for the heuristic solution, $Z^H(I)$, is also not available, we introduce an (attainable) inferior surrogate heuristic, H', in which $Z^{H'}(I)$ has an explicit simple form. We let $f_H(I)$, which we call the surrogate worst-case performance ratio function for heuristic H, equal to $\frac{Z^{H'}(I)}{Z^{S*}(I)}$. This function serves as a lower bound for the worst-case performance ratio of the heuristic given I. If the bound is tight, a worst case performance ratio can be obtained by taking the infimum of the function $f_H(I)$ over all problem instances. Since a problem instance I can be represented as a point in R^k, the derivation of a worst case performance ratio boils down to the minimization of f_H over an admissible region $S_H \subset R^k$. If the admissible region S_H is difficult to characterize, to simplify the analysis, we further introduce a relaxed admissible region S'_H, which is a superset of the admissible region. We summarize our approach as follows:

$$\begin{aligned}\eta_H &\equiv \inf_I\{\frac{Z^H(I)}{Z^*(I)}\}\\ &\geq \inf_I\{\frac{Z^{H'}(I)}{Z^{S*}(I)}\}\\ &\equiv \inf_I\{f_H(I)\}\\ &\equiv \inf\{f_H(x):\ x\in S_H\}\\ &\geq \inf\{f_H(x):\ x\in S'_H\}.\end{aligned}$$

We note that this optimization approach involves two steps of relaxations: the use of surrogate algorithms and the use of the relaxed admissible region. The bound may or may not be tight depending on how we choose the surrogate algorithms and the relaxed admissible region.

To illustrate our approach, we analyze a flexible 0-1 knapsack problem, denoted by FKP, which can be formulated as follows:

(FKP)

$$\begin{aligned}&\max \sum_{j=1}^{n} p_j x_j + c_1 y_1 - c_2 y_2\\ \text{subject to:}\ &\sum_{j=1}^{n} w_j x_j \leq b - y_1 + y_2\\ &x_j = 0, 1,\ j = 0, 1, \ldots, n\\ &y_1, y_2 \geq 0,\end{aligned}$$

where c_2 is the unit cost to expand capacity and c_1 is the unit salvage value of remaining capacity. Note that if b represents "capital", then c_2 and c_1 can be interpreted as the borrowing and the lending rates of "capital", respectively. We assume b, p_j, w_j are positive rational numbers satisfying $w_j \leq b$ for all j and $c_2 > c_1 > 0$. In section 4 we will discuss the effect of relaxing the assumption that $w_j \leq b$ for all j. Note that the constraint $y_1 y_2 = 0$ is dropped in FKP since it is automatically satisfied in the optimal solution.

FKP is a variation of the 0-1 knapsack problem. The 0-1 knapsack problem, denoted by KP, can be formulated as follows:

$$\text{maximize } \{\sum_{j=1}^{n} p_j x_j : \sum_{j=1}^{n} w_j x_j \leq b,\ x_j = 0,\ 1,\quad j = 1,\ \ldots,\ n\}.$$

If $x_j = 0, 1, \ldots, u_j$, $j = 1, 2, \ldots, n$, then the problem is the bounded knapsack problem [6], denoted by BKP. If x_j, $j = 1, 2, \ldots, n$ can take nonnegative integers, then the problem is the unbounded knapsack problem [6], denoted by UKP. If $p_j = w_j$, $j = 1, 2, \ldots, n$, then the problem is the subset sum problem [6], denoted by SSP. We note that KP is a special case of BKP (by letting $u_j = 1\ \forall j$); UKP is a special case of BKP (by letting $u_j = \lfloor \frac{b}{w_j} \rfloor\ \forall j$); SSP is a special case of KP. We also note that both BKP and UKP can be transformed into KP by adding more binary variables. It can be shown [5] that UKP is NP-hard by transformation from SSP.

We can define the "flexible versions" of BKP, UKP and SSP, denoted by FBKP, FUKP and FSSP, respectively, in a similar manner. We also note that FKP, FBKP, FUKP and FSSP are all NP-hard since each contains as a special case (when $c_1 = 0$ and $c_2 = \infty$) KP, BKP, UKP and SSP, respectively, each of which is known to be NP-hard [3]. We then focus on approximate algorithms for FKP.

The remainder of this paper is organized as follows. In section 2, each of two basic heuristics (High Ratio First and High Value First) for the standard 0-1 knapsack problem is customized and expanded to two versions by utilizing the "soft" constraint in FKP. We derive the worst case performance ratios for each of these four individual heuristics and provide the background for later analysis. In section 3, we analyze the worst case behaviors of three different compound heuristics which are formed by combining two (or more) of the individual heuristics developed in section 2. We prove analytically that the worst case performance of each of the three compound heuristics is superior to the worst case performance of the individual heuristics. In section 4, data-dependent worst case performance ratios for the individual and compound heuristics are derived in terms of a summary statistic of the problem input data. In section 5, we extend our results to other types of flexible knapsack problems. We conclude with section 6.

2. Four Basic Heuristics for FKP

In this section, we present four individual heuristics and provide the background for later analysis. Two basic heuristics for the standard 0-1 knapsack problem are the high ratio first (HRF) and high value first (HVF) heuristics (Fisher 1980). HRF (known as the Greedy Algorithm) selects items with higher ratios of $\frac{p_j}{w_j}$ first, while HVF selects items with higher values of p_j first. Since the capacity constraint in FKP is "soft", we will consider two versions (R1 vs. R2 or V1 vs. V2) of HRF or HVF, depending on whether we underuse or overuse the prespecified capacity b, respectively. Formal descriptions of these heuristics are as follows:

R1: 1. Reindex the items so that $\frac{p_1}{w_1} \geq \frac{p_2}{w_2} \geq \ldots \geq \frac{p_n}{w_n}$. Set $j = 1$ and $B = b$.
2. If $w_j \leq B$ then set $x_j = 1$ and $B = B - w_j$; otherwise, set $x_j = 0$.
3. Stop if $j = n$ or $B = 0$; otherwise, set $j = j + 1$ and go to step 2.

R2: 1. Reindex the items so that $\frac{p_1}{w_1} \geq \frac{p_2}{w_2} \geq \ldots \geq \frac{p_n}{w_n}$. Set $j = 1$ and $B = b$.
2. If $w_j \leq B$ then set $x_j = 1$ and $B = B - w_j$; otherwise, set $x_j = 1$ and stop.
3. Set $j = j + 1$ and go to step 2.

V1: 1. Reindex the items so that $p_{[1]} \geq p_{[2]} \geq \ldots \geq p_{[n]}$. Set $j = [1]$ and $B = b$.
2. If $w_{[j]} \leq B$ then set $x_{[j]} = 1$ and $B = B - w_{[j]}$; otherwise, set $x_{[j]} = 0$.
3. Stop if $j = [n]$ or $B = 0$; otherwise, set $[j] = [j + 1]$ and go to step 2.

V2: 1. Reindex the items so that $p_{[1]} \geq p_{[2]} \geq \ldots \geq p_{[n]}$. Set $j = [1]$ and $B = b$.
2. If $w_{[j]} \leq B$ then set $x_{[j]} = 1$ and $B = B - w_{[j]}$;
otherwise, set $x_{[j]} = 1$ and stop.
3. Set $[j] = [j + 1]$ and go to step 2.

Note that both R2 and V2 can yield a negative solution. In this case, we assume that we choose the do-nothing option and thus obtain a solution with value zero. We will assume for the remainder of this paper that the items have been indexed so that $r_1 \geq r_2 \geq \ldots \geq r_n$, where $r_i = \frac{p_i}{w_i}, i = 1, 2, \ldots, n$, and $p_{[1]} \geq p_{[2]} \geq \ldots \geq p_{[n]}$. Also, we assume that $c_2 > r_i > c_1$, for all i, since it is optimal to include all items with $r_i \geq c_2$ and to exclude all items with $r_i \leq c_1$.

For R1 and R2, we define $W_k = \sum_{i=1}^{k} w_i$, $P_k = \sum_{i=1}^{k} p_i$, and $R_k = \frac{P_k}{W_k}$, $k = 1,2,\ldots,n$; thus, W_k, P_k, R_k, are the cumulative (up to k) weights, profits, and average ratios (profit densities), respectively. We define the *marginal item* $m+1$ as the item such that $W_m \leq b$ and $W_{m+1} > b$. Note that R2 picks up the first $m+1$ items while R1 picks up the first m items plus some items behind item m. Similarly, for V1 and V2, we define $W_{[k]} = \sum_{i=1}^{k} w_{[i]}$, $P_{[k]} = \sum_{i=1}^{k} p_{[i]}$, and $R_{[k]} = \frac{P_{[k]}}{W_{[k]}}$, $k = 1,2,\ldots,n$; and we define the *marginal item* $l+1$ as the item such that $W_{[l]} \leq b$ and $W_{[l+1]} > b$. V2 picks up the first $l+1$ items while V1 picks up the first l items plus some items behind item l.

We let Z^{LP} be an upper bound obtained by relaxing the integer constraints. It is instructive to mention here that we will use Z^{LP} as an upper bound on Z^* since it can be explicitly expressed and is attainable (thus the tightness of the bound in the worst case).

Since R2 and V2 pick up the first $m+1$ and $l+1$ items, respectively, expressions of simple forms for Z^{R2} and Z^{V2} are available. However, since R1 and V1 pick up, in addition to the first m and l items, respectively, some unknown items behind items m and l, respectively, no expressions of simple forms for Z^{R1} and Z^{V1} are available. Hence, to facilitate the derivations of the worst case performance ratios for R1, V1 and their associated compound heuristics, we use one surrogate heuristic, denoted by R1L, for R1 and two surrogate heuristics, denoted by V1L1 and V1L2 respectively, for V1. For R1L, we pick up the first m items only; thus Z^{R1L} is a (attainable) lower bound for Z^{R1}. For V1L1 and V1L2, we only pick up the first l items and item $m+1$, respectively, and thus Z^{V1L1} and Z^{V1L2} are (attainable) lower bounds for Z^{V1}.

The following are expressions for Z^{R1L}, Z^{R2}, Z^{V1L1}, Z^{V1L2}, Z^{V2}, and Z^{LP} in terms of the quantities defined above:

$$Z^{R1} \geq Z^{R1L} \equiv R_m W_m + c_1(b - W_m) \tag{1}$$

$$Z^{R2} = R_m W_m + r_{m+1} w_{m+1} - c_2(W_{m+1} - b) \tag{2}$$

$$Z^{V1} \geq Z^{V1L1} \equiv R_{[l]} W_{[l]} + c_1(b - W_{[l]}) \tag{3}$$

$$Z^{V1} \geq Z^{V1L2} \equiv r_{m+1} w_{m+1} + c_1(b - w_{m+1}) \tag{4}$$

$$Z^{V2} = R_{[l]} W_{[l]} + r_{[l+1]} w_{[l+1]} - c_2(W_{[l+1]} - b) \tag{5}$$

$$Z^{LP} = R_m W_m + r_{m+1}(b - W_m). \tag{6}$$

Since an explicit form for Z^* is not available, we introduce an attainable upper bound Z^{LP} as a substitute. Since an explicit form for Z^{R1} or Z^{V1} is also not available, we use R1L as a substitute for R1 and V1L1 or V1L2 for V1.

Thus, we define

$$\begin{aligned}
&f_1(R_m, r_{m+1}, W_m, w_{m+1}) = Z^{R1L}/Z^{LP}, \\
&f_2(R_m, r_{m+1}, W_m, w_{m+1}) = Z^{R2}/Z^{LP}, \\
&f_{31}(R_{[l]}, W_{[l]}, R_m, r_{m+1}, W_m) = Z^{V1L1}/Z^{LP}, \\
&f_{32}(R_m, r_{m+1}, W_m, w_{m+1}) = Z^{V1L2}/Z^{LP}, \\
\text{and}\quad &f_4(R_{[l]}, r_{[l+1]}, W_{[l]}, w_{[l+1]}, R_m, r_{m+1}, W_m) = Z^{V2}/Z^{LP}.
\end{aligned}$$

Note that f_1, f_2, and f_4 are lower bounds for η_{R1}, η_{R2}, and η_{V2}, respectively, while f_{31} and f_{32} are lower bounds for η_{V1}.

We find it convenient to express our "(surrogate) worst case performance ratio functions" f_1, f_2 and f_{32} in terms of R_m, r_{m+1}, W_m, w_{m+1}; f_{31} in terms of $R_{[l]}$, $W_{[l]}$, R_m, r_{m+1}, W_m; and f_4 in terms of $R_{[l]}$, $r_{[l+1]}$, $W_{[l]}$, $w_{[l+1]}$, R_m, r_{m+1}, W_m.
Lemma 1, which follows, characterizes the behaviors of these functions in terms of their arguments. We will use lemma 1 to derive worst case performance ratios for each of the four individual heuristics (lemma 2), as well as for three different compound heuristics (in section 3).

Lemma 1.

(1) $f_1(R_m, r_{m+1}, W_m, w_{m+1})$ is increasing in R_m, decreasing in r_{m+1}, increasing in W_m, and independent of w_{m+1}.

(2) $f_2(R_m, r_{m+1}, W_m, w_{m+1})$ is increasing in R_m, increasing in r_{m+1}, decreasing in W_m, and decreasing in w_{m+1}.

(3) $f_{31}(R_{[l]}, W_{[l]}, R_m, r_{m+1}, W_m)$ is increasing in $R_{[l]}$, increasing in $W_{[l]}$, decreasing in R_m, decreasing in r_{m+1}, and decreasing in W_m.

(4) $f_{32}(R_m, r_{m+1}, W_m, w_{m+1})$ is decreasing in R_m, increasing in r_{m+1}, decreasing in W_m, and increasing in w_{m+1}.

(5) $f_4(R_{[l]}, r_{[l+1]}, W_{[l]}, w_{[l+1]}, R_m, r_{m+1}, W_m)$ is increasing in $R_{[l]}$, increasing in $r_{[l+1]}$, decreasing in $W_{[l]}$, and decreasing in $w_{[l+1]}$. Furthermore,

If $c_1 > (1/2)c_2$, then $f_4(R_{[l]}, r_{[l+1]}, W_{[l]}, w_{[l+1]}, R_m, r_{m+1}, W_m)$ is decreasing in R_m, decreasing in r_{m+1}, and decreasing in W_m.

If $c_1 \leq (1/2)c_2$, then $f_4(R_{[l]}, r_{[l+1]}, W_{[l]}, w_{[l+1]}, R_m, r_{m+1}, W_m)$ is increasing in R_m, increasing in r_{m+1}, and increasing in W_m.

Proof. The result can be verified by examining partial derivatives. ◇

Let $S_1 \subset R^4$, $S_2 \subset R^4$, $S_3 \subset R^5$ and $S_4 \subset R^7$ be the admissible regions characterized by R1L plus LP, R2 plus LP, V1L1 plus LP, and V2 plus LP, respectively. To simplify the analysis, it is useful to introduce a relaxed surrogate region S_i' for S_i, $i = 1, 2, 3, 4$, where

$$S_1' = S_2' = \{c_2 > R_m \geq r_{m+1} > c_1,\ b \geq W_m, w_{m+1} > 0,\ W_m + w_{m+1} > b\},$$

$$S_3' = \{c_2 > R_{[l]} > c_1,\ b > W_{[l]} > 0,\ c_2 > R_m \geq r_{m+1} > c_1,\ b \geq W_m, w_{m+1} > 0\},$$

$$S_4' = \{b > W_{[l]}, w_{[l+1]} > 0,\ W_{[l]} + w_{[l+1]} > 0,\ c_2 > R_m \geq r_{m+1} > c_1,\ b \geq W_m, w_{m+1} > 0\}.$$

Since R1L plus LP, R2 plus LP, V1L1 plus LP, and V2 plus LP satisfy the constraints specified by S_1', S_2', S_3', and S_4', respectively, we have $S_i \subset S_i'$, $i = 1,2,3,4$. Furthermore, since the infimum of f_i, $(i = 1,2,3,4)$ over S_i is a lower bound for the worst case performance ratio and since $S_i \subset S_i'$, the infimum of f_i over S_i' is also a lower bound. In fact, it can be shown that $S_1 = S_2 = S_1' = S_2'$. Also, note that the introduction of a relaxed surrogate region can yield a gap in the bound, but, as we will show, the regions S_i' do not yield any gap. In the proof of lemma 2, we first establish a lower bound for the worst case performance ratio of each of the four heuristics and then show that this bound is tight by constructing an example.

Lemma 2.

(1) $\eta_{R1} = \frac{c_1}{c_2}$,
(2) $\eta_{R2} = \max\{2 - \frac{c_2}{c_1},\ 0\}$,
(3) $\eta_{V1} = \frac{c_1}{c_2}$,
(4) $\eta_{V2} = \begin{cases} \frac{2c_1}{c_2} - 1, & \text{if } c_1 > (1/2)c_2; \\ 0, & \text{if } c_1 \le (1/2)c_2. \end{cases}$

Proof. (1) By applying lemma 1 iteratively for each argument, we have

$$\eta_{R1} \ge \inf\{f_1(R_m, r_{m+1}, W_m, w_{m+1}) :\ S_1\} = f_1(c_2, c_2, 0, b) = \frac{c_1}{c_2}.$$

To show that the bound is tight, consider the example: $n = 2$, $p_1 = c_2 - \epsilon_1$, $w_1 = 1$, $p_2 = (c_2 - \epsilon_2)K$, $w_2 = K$, $b = K$, where $\epsilon_2 > \epsilon_1 > 0$. Then $Z^{R1} = c_2 - \epsilon_1 + (K-1)c_1$ and $Z^* = K(c_2 - \epsilon_2)$. Hence, $\frac{Z^{R1}}{Z^*} \to \frac{c_1}{c_2}$ as $K \to \infty$ and $\epsilon_2 \to 0$.

Similarly, we can obtain the results for (2), (3) and (4). ◇

3. Analysis of Compound Heuristics for FKP

In this section, two (or more) heuristics which can complement one another in the sense of worst case performance are combined to form a better compound heuristic [2] [4], [7]. In particular, we consider three such heuristics by combining (1) R1 and R2, (2) R1 and V1, and (3) R1, R2 and V1. The viability of such a heuristic depends on the extent to which the two (or more) heuristics employed do not share their worst case input.

Since, for a compound heuristic, we select the best solution from the two (or more) separately generated solutions, the derivation of a worst case performance ratio for a compound heuristic (in the case of a maximization problem) involves solving a minimax problem. The following lemma allows us, under some conditions, to convert a minimax problem into a minimization problem, and thus is useful in the subsequent analysis. Lemma 3 says that if two continuous functions are monotonic in opposite directions for each argument and neither dominates the other over a connected feasible region S, then a minimax problem over S can be converted to a minimization problem over a more restricted region S_0, where S_0 is that subset of the domain over which the two functions have equal value.

Lemma 3. Let

$$h(y_1, \ldots, y_n) = \max\{h_1(y_1, \ldots, y_n), h_2(y_1, \ldots, y_n)\},$$

where there exist at least one $(y_1, \ldots, y_n) \in S \subset R^n$ and at least one $(\bar{y_1}, \ldots, \bar{y_n}) \in S \subset R^n$ such that

$$h(y_1, \ldots, y_n) > h_1(y_1, \ldots, y_n) \quad \text{and} \quad h(\bar{y_1}, \ldots, \bar{y_n}) > h_2(\bar{y_1}, \ldots, \bar{y_n}).$$

If a continuous function $h_1(y_1, \ldots, y_n)$ is decreasing (increasing) in y_i, $i = 1, \ldots, n$, and if a continuous function $h_2(y_1, \ldots, y_n)$ is increasing (decreasing) in y_i, $i = 1, \ldots, n$, then

$$\begin{aligned} A &\equiv \inf\{\max\{h_1(y_1, \ldots, y_n),\ h_2(y_1, \ldots, y_n)\} :\ (y_1, \ldots, y_n) \in S \subset R^n\} \\ &= \inf\{h_1(y_1, \ldots, y_n) :\ (y_1, \ldots, y_n) \in S_0\} \equiv B, \end{aligned}$$

where S is connected and

$S_0 = \{(y_1, \ldots, y_n) : h_1(y_1, \ldots, y_n) = h_2(y_1, \ldots, y_n),\ S \subset R^n\}$ is not empty.

Proof. (contradiction) We want to show that $A = B$. It is clear that $A \leq B$ since $S_0 \subset S$. Therefore, we now establish that $A \geq B$. Given any $y^0 = (y_1^0, \ldots, y_n^0) \in S$, we need to show that there exists a $y' = (y_1', \ldots, y_n') \in S_0$ such that $h(y^0) \geq h(y')$. Fixing y^0, without loss of generality, we assume that $h_1(y^0) > h_2(y^0)$. Without loss of generality, we also assume that $h_1(y)$ is increasing in y_i for $i \in N_1 = \{1, \ldots, m\}$ and $h_1(y)$ decreasing in y_i for $i \in N_2 = \{m+1, \ldots, n\}$; thus, $h_2(y)$ is decreasing in y_i for $i \in N_1$ and increasing in y_i for $i \in N_2$. Let $y^0 = (y_a^0, y_b^0)$, where $y_a^0 = (y_1^0, \ldots, y_m^0)$ and $y_b^0 = (y_{m+1}^0, \ldots, y_n^0)$. Define $S_1 = \{(y_a, y_b) : y_a \leq y_a^0,\ y_b \geq y_b^0,\ (y_a, y_b) \in S\}$ and $S_2 = S - S_1$. Since S_0 is not empty, either $y' \in S_1$ or $y' \in S_2$. If $y' \in S_1$: since $h(y^0) \geq h(y)$ for all $y \in S_1$, we have $h(y^0) \geq h(y')$. If $y' \in S_2$ but $y' \notin S_1$: since, by assumption, we know that $h_1(y') = h_2(y')$, there exists a $y'' = (y_a'', y_b'') \in S_1$ such that $h_1(y'') < h_2(y'')$. Since $h_1(y^0) > h_2(y^0)$ and $h_1(y'') < h_2(y'')$, y^0, $y'' \in S_1$, and since S_1 is connected and h_1 and h_2 are continuous, there exists $y''' \in S_1$ such that $h_1(y''') = h_2(y''')$, which is a contradiction. This completes our proof. ◇

We now use lemmas 1 and 3 to analyze three different compound heuristics.

Recall that R1 consistently underuses the prespecified capacity and R2 consistently overuses the prespecified capacity. An improved heuristic (MR) would use R1 and R2 separately to generate two solutions and then select the better one. We define

$$g_1(R_m, r_{m+1}, W_m, w_{m+1}) = \max\{f_1(R_m, r_{m+1}, W_m, w_{m+1}), f_2(R_m, r_{m+1}, W_m, w_{m+1})\}.$$

Also, we let

$$g_1^* = g_1(R_m^*, r_{m+1}^*, W_m^*, w_{m+1}^*) = \inf\ \{g_1(R_m, r_{m+1}, W_m, w_{m+1}) : S_1'\}.$$

Since R1L plus LP and R2 plus LP satisfy the constraints specified by S_1',

$$\eta_{MR} \geq g_1^* = g_1(R_m^*, r_{m+1}^*, W_m^*, w_{m+1}^*).$$

Theorem 1:

$$\eta_{MR} = \frac{2(\sqrt{c_1 c_2} - c_1)}{(c_2 - c_1)}.$$

Proof. By lemma 1, since g_1 is decreasing in w_{m+1} and $w_{m+1} \leq b$, thus $w_{m+1}^* = b$; similarly, g_1 is increasing in R_m and $R_m \geq r_{m+1}$, so $R_m^* = r_{m+1}$. Thus, only W_m^* and r_{m+1}^* remain to be determined. Applying lemma 3, we can deduce $W_m^* = \frac{b(r_{m+1} - c_1)}{c_2 - c_1}$ from the constraint $f_1 = f_2$. Now, we have

$$g_1^* = \inf\ \{\frac{r_{m+1}^2 - 2r_{m+1}c_1 + c_1 c_2}{r_{m+1}(c_2 - c_1)} : c_2 > r_{m+1} > c_1\},$$

which is minimized at $r_{m+1}^* = \sqrt{c_1 c_2}$. Thus, $\eta_{MR} \geq g_1^* = \frac{2(\sqrt{c_1 c_2} - c_1)}{c_2 - c_1}$.

To show that the bound is tight, consider the example: $n = 2$, $p_1 = \sqrt{c_1 c_2}\frac{\sqrt{c_1 c_2} - c_1}{c_2 - c_1}$, $w_1 = \frac{(\sqrt{c_1 c_2} - c_1)}{c_2 - c_1}$, $p_2 = \sqrt{c_1 c_2}$, $w_2 = 1$, $b = 1$. Then $Z^{MR} = 2\sqrt{c_1 c_2}\frac{\sqrt{c_1 c_2} - c_1}{c_2 - c_1}$ and $Z^* = \sqrt{c_1 c_2}$. Hence, $\frac{Z^{MR}}{Z^*} = \frac{2(\sqrt{c_1 c_2} - c_1)}{c_2 - c_1}$. ◇

It is straightforward to show that $\eta_{MR} > \max\{\eta_{R1}, \ \eta_{R2}\}$.

We now consider a second compound heuristic, RV11, formed using R1 and V1, and define

$$g_2(R_m, r_{m+1}, W_m, w_{m+1}) = \max\{f_1(R_m, r_{m+1}, W_m, w_{m+1}), f_{32}(R_m, r_{m+1}, W_m, w_{m+1})\}.$$

Also, we let

$$g_2^* = g_2(R_m^*, r_{m+1}^*, W_m^*, w_{m+1}^*) = \inf\ \{g_2(R_m, r_{m+1}, W_m, w_{m+1}) \ : \ S_1'\}.$$

Since R1L plus LP and V1L2 plus LP satisfy the constraints specified by S_1',

$$\eta_{RV11} \geq \ g_2^* = g_2(R_m^*, r_{m+1}^*, W_m^*, w_{m+1}^*).$$

Then:

Theorem 2:

$$\eta_{RV11} = \frac{1}{2}(1 + \frac{c_1}{c_2}).$$

Proof. By lemma 1, g_2 is increasing in w_{m+1}. Since $W_m + w_{m+1} > b$, we know that $w_{m+1}^* = b - W_m$. Therefore, we still have W_m^*, R_m^* and r_{m+1}^* remained to be determined. Applying lemma 3, we can deduce $W_m^* = \frac{b(r_{m+1} - c_1)}{R_m + r_{m+1} - 2c_1}$ from the constraint $f_1 = f_{32}$. Now, we have

$$g_2^* = \inf\ \{\frac{R_m r_{m+1} - c_1^2}{2R_m r_{m+1} - c_1(R_m + r_{m+1})} \ : \ c_2 > R_m \geq r_{m+1} > c_1\}.$$

This expression is decreasing in R_m, as can be verified by examining its partial derivative with respect to R_m. Thus, we have $R_m^* = c_2$ and hence

$$g_2^* = \inf\ \{\frac{c_2 r_{m+1} - c_1^2}{2c_2 r_{m+1} - c_1(c_2 + r_{m+1})} \ : \ c_2 > r_{m+1} > c_1\}.$$

This expression is decreasing in r_{m+1}, as can be verified by examining its partial derivative with respect to r_{m+1}. Thus, we have $r_{m+1}^* = c_2$ and hence $g_2^* = \frac{1}{2}(1 + \frac{c_1}{c_2})$. Thus, $\eta_{RV11} \geq g_2^* = \frac{1}{2}(1 + \frac{c_1}{c_2})$.

To show that the bound is tight, consider the example: $n = n_0 + 3$, $p_i = c_1 + \epsilon_1$, $w_i = \frac{c_1}{c_2}$, $i = 1, \ldots, n_0$, $p_{n_0+1} = c_2 e$, $w_{n_0+1} = e + \epsilon_2$, $p_{n_0+2} = c_2$, $w_{n_0+2} = 1 + \epsilon_3$, $p_{n_0+3} = c_1 + \epsilon_4$, $w_{n_0+3} = 1 - \epsilon_5$, $b = 2$, where $n_0 = \lfloor \frac{c_2}{c_1} \rfloor$, $\epsilon_4 > \epsilon_1 > 0$, $\epsilon_5 > \epsilon_3 > \epsilon_2 > 0$, $e = 1 - n_0 \frac{c_1}{c_2}$, $\epsilon_3 + \epsilon_2 > e > \epsilon_3$. Then $Z^{R1} = c_2 + c_1 + \epsilon_a$, $Z^{V1} = c_2 + c_1 + \epsilon_4$, and $Z^* = 2c_2 + \epsilon_b$, where $\epsilon_a = n_0 \epsilon_1 + \epsilon_4$ and $\epsilon_b = n_0 \epsilon_1$. Hence, $\frac{Z^{RV11}}{Z^*} \rightarrow \frac{1}{2}(1 + \frac{c_1}{c_2})$ as $\epsilon_i \rightarrow 0, i = 1, 2, 3, 4, 5$. ◇

It is straightforward to show that $\eta_{RV11} > \max\{\eta_{R1}, \ \eta_{V1}\}$. Also, note that neither RV11 nor MR dominates the other in the sense of worst case performance: MR is better (in the worst case sense) than RV11 when $\frac{c_1}{c_2} > 3 - 2\sqrt{2}$, while RV11 is better than MR when $\frac{c_1}{c_2} < 3 - 2\sqrt{2}$.

We now consider RRV, the compound heuristic of R1, R2 and V1. We define

$$g_3(R_m, r_{m+1}, W_m, w_{m+1}) = \max\{f_1(R_m, r_{m+1}, W_m, w_{m+1}),\\ f_2(R_m, r_{m+1}, W_m, w_{m+1}), f_{32}(R_m, r_{m+1}, W_m, w_{m+1})\}.$$

Also, we let $g_3^* = g_3(R_m^*, r_{m+1}^*, W_m^*, w_{m+1}^*) = \inf\{g_3(R_m, r_{m+1}, W_m, w_{m+1}) : S_1'\}$. Since R1L plus LP, R2 plus LP and V1L2 plus LP satisfy the constraints specified by S_1', $\eta_{RRV} \geq g_3^* = g_3(R_m^*, r_{m+1}^*, W_m^*, w_{m+1}^*)$. We let $g_{31}^* = \inf\{f_1 : f_1 \geq f_2, f_1 \geq f_{32}, S_1'\}$, $g_{32}^* = \inf\{f_2 : f_2 \geq f_1, f_2 \geq f_{32}, S_1'\}$, $g_{33}^* = \inf\{f_{32} : f_{32} \geq f_1, f_{32} \geq f_2, S_1'\}$. Then, $g_3^* = \min\{g_{31}^*, g_{32}^*, g_{33}^*\}$.

The following lemma is needed in the proof of theorem 3.

Lemma 4: For i=1,2, $g_{3i}^* = \inf\{f_i : f_1 = f_2 = f_{32}, S_1'\}$, and $i = 3$,

$$g_{3i}^* = \inf\{f_{32} : f_1 = f_2 = f_{32}, S_1'\}, \text{ i.e., } g_3^* = g_{31}^* = g_{32}^* = g_{33}^*.$$

Proof. For $i = 1$, we first note that, from the constraints $f_1 \geq f_2$ and $f_1 \geq f_{32}$ in g_{31}^*, we can deduce

$$W_m \geq b - \frac{c_2 - r_{m+1}}{c_2 - c_1} w_{m+1} \equiv h_1(w_{m+1}) \text{ and } W_m \geq \frac{r_{m+1} - c_1}{R_m - c_1} w_{m+1} \equiv h_2(w_{m+1}).$$

Since, by lemma 1(1), f_1 is increasing in W_m, g_{31}^* achieves the infimum when $W_m = \min\{h_1(w_{m+1}), h_2(w_{m+1})\}$. Since h_1 is decreasing in w_{m+1} and h_2 is increasing in w_{m+1}, we know by lemma 3 that g_{31}^* achieves the infimum when $W_m = h_1(w_{m+1}) = h_2(w_{m+1})$, which are equivalent to the constraints $f_1 = f_2 = f_{32}$. Thus, $g_{31}^* = \inf\{f_1 : f_1 = f_2 = f_{32}, S_1'\}$.

A similar analysis proves the result for $i = 2, 3$. ◇

We now use lemma 4 to establish the worst case performance ratio for RRV.

Theorem 3:

$$\eta_{RRV} = \frac{(c_2 - c_1)}{2(c_2 - \sqrt{c_1 c_2})}$$

Proof. Since $g_3^* = \min\{g_{31}^*, g_{32}^*, g_{33}^*\}$, by lemma 4, $g_3^* = g_{31}^* = \inf\{f_1 : f_1 = f_2, f_1 = f_{32}, S_1'\}$. From the constraints $f_1 = f_2 = f_{32}$, we can obtain $w_{m+1} = (R_m - c_1)A$ and $W_m = (r_{m+1} - c_1)A$, where
$A = \frac{(c_2 - c_1)b}{(r_{m+1} - c_1)(c_2 - c_1) + (R_m - c_1)(c_2 - r_{m+1})}$. Hence, we can express

$$f_1 = \frac{R_m(c_2 r_{m+1} + c_1^2 - 2c_1 r_{m+1}) - c_1^2(c_2 - r_{m+1})}{R_m(2c_2 r_{m+1} + c_1^2 - c_1 r_{m+1} - c_1 c_2 - r_{m+1}^2) - c_1 r_{m+1}(c_2 - r_{m+1})}.$$

Now, f_1 is decreasing in R_m, as can be verified by examining its partial derivative with respect to R_m. Thus, we have $R_m^* = c_2$ and we can then express

$$f_1 = \frac{r_{m+1}(c_2 - c_1)}{-r_{m+1}^2 + 2c_2 r_{m+1} - c_1 c_2},$$

which is minimized at $r_{m+1}^* = \sqrt{c_1 c_2}$ (which satisfies $c_2 > r_{m+1}^* > c_1$), so:

$g_3^* = \frac{(c_2-c_1)}{2(c_2-\sqrt{c_1c_2})}$. Thus, $\eta_{RRV} \geq g_3^* = \frac{(c_2-c_1)}{2(c_2-\sqrt{c_1c_2})}$.

To show that the bound is tight, consider the example:
$n = 3$, $p_1 = c_2\frac{\sqrt{c_1c_2}-c_1}{c_2-c_1}$, $w_1 = \frac{(\sqrt{c_1c_2}-c_1)}{c_2-c_1}$, $p_2 = \sqrt{c_1c_2}+\epsilon$, $w_2 = 1$, $p_3 = \sqrt{c_1c_2}\frac{\sqrt{c_1c_2}-c_1}{c_2-c_1}$, $w_3 = \frac{(\sqrt{c_1c_2}-c_1)}{c_2-c_1}$, $b = 1$, where $\epsilon > 0$. Then $Z^{RRV} = \sqrt{c_1c_2}+\epsilon$ and $Z^* = 2\sqrt{c_1c_2}\frac{c_2-\sqrt{c_1c_2}}{c_2-c_1}$. Hence, as $\epsilon \to 0$, $\frac{Z^{RRV}}{Z^*} = \frac{(c_2-c_1)}{2(c_2-\sqrt{c_1c_2})}$. ◇

It is straightforward to show that $\eta_{RRV} > \max\{\eta_{R1}, \eta_{R2}, \eta_{V1}\}$.

Although each of the three compound heuristics considered above consists of individual heuristics which complement one another in the case of worst case performance, we note that, often, individual heuristics may not complement one another in their worst case performance. Generally speaking, intuition plays an important role in the selection of complementary heuristics for a compound heuristic. However, intuition can be wrong. For example, since V1 consistently underuses and V2 consistently overuses the prespecified capacity, b, one might expect that heuristics V1 and V2 are complementary in their worst case performance. Let MV denote the compound heuristic of V1 and V2. The following example shows that $\eta_{MV} = \eta_{V1} = \max\{\eta_{V1}, \eta_{V2}\}$, i.e., V1 and V2 do not complement each other: $n = K+1$, $p_1 = c_1K+1$, $w_1 = K$, $p_i = c_2-\epsilon$, $w_i = 1, i = 2, \ldots, K+1$, $b = K$, where $\epsilon > 0$. Then $Z^{MV} = c_1K + 1$ and $Z^* = (c_2 - \epsilon)K$. Hence, $\frac{Z^{MV}}{Z^*} \to \frac{c_1}{c_2}$ as $K \to \infty$ and $\epsilon \to 0$.

4. Data-Dependent Bounds

In this section, we derive the worst case performance ratios for the seven heuristics considered above in terms of a summary statistic α of the problem input data, where $\alpha = \min_j\{\frac{b}{w_j}\}$. The integral part of α represents the minimum number of items we can put into the knapsack. Note that under our assumption that $w_j \leq b\ \forall j$, we have $\alpha \geq 1$. We also assume that $\alpha < n$; otherwise the problem trivially reduces to one where we can pick all the items.

As a function of α, the worst case performance ratios for the four basic individual heuristics are derived in lemma 5, which follows. The proof is straightforward if we follow the proof in lemma 2.

Lemma 5.

(1) $\eta_{R1} = \frac{1}{\alpha}(\alpha - 1 + \frac{c_1}{c_2})$,

(2) $\eta_{R2} = \max\{1 + \frac{1}{\alpha}(1 - \frac{c_2}{c_1}),\ 0\}$

(3) $\eta_{V1} = \frac{c_1}{c_2}$,

(4) $\eta_{V2} = \begin{cases} \frac{c_1}{c_2} + \frac{1}{\alpha}(\frac{c_1}{c_2} - 1), & \text{if } c_1 > (\frac{1}{\alpha+1})c_2; \\ 0, & \text{if } c_1 \leq (\frac{1}{\alpha+1})c_2. \end{cases}$

Proof. The proof is similar to that of lemma 2 except that the constraint set S_i' is changed. We now have $\frac{b}{\alpha} > w_{m+1} > 0$ and/or $\frac{b}{\alpha} > w_{[l+1]} > 0$ instead of $b > w_{m+1} > 0$ and/or $b > w_{[l+1]} > 0$ in the proof of lemma 2. The proof of lemma 5 follows immediately. ◇

Theorem 4 gives the worst case performance ratios in terms of α for the compound heuristics MR, RV11, and RRV.

Theorem 4.

(1)
$$\eta_{MR} = \begin{cases} \frac{2(\sqrt{c_1c_2}-c_1)}{(c_2-c_1)}, & \text{if } \alpha \le 1+\sqrt{\frac{c_1}{c_2}}; \\ \frac{1}{c_2-c_1}\left(\frac{(\alpha-1)c_2+c_1}{\alpha} + \frac{\alpha c_1 c_2}{(\alpha-1)c_2+c_1} - 2c_1\right), & \text{if } \alpha > 1+\sqrt{\frac{c_1}{c_2}}. \end{cases}$$

(2)
$$\eta_{RV11} = \begin{cases} \frac{1}{2}(1+\frac{c_1}{c_2}), & \text{if } \alpha < 2; \\ \frac{1}{\alpha}(\alpha - 1 + \frac{c_1}{c_2}), & \text{if } \alpha \ge 2. \end{cases}$$

(3) If $\alpha < 2$, then

$$\eta_{RRV} = \begin{cases} \frac{(c_2-c_1)c_2-(\alpha-1)(c_2-c_1)^2}{c_2^2-(\alpha-1)^2(c_2-c_1)^2-c_1c_2}, & \text{if } \delta(\alpha) < 0; \\ \frac{(c_2-c_1)\sqrt{\alpha(\alpha-1)c_1^2+\alpha c_1c_2}}{2(c_2+(\alpha-1)c_2)\sqrt{\alpha(\alpha-1)c_1^2+\alpha c_1c_2}-2\alpha c_1c_2-(\alpha-1)(2\alpha-1)c_1^2}, & \text{if } \delta(\alpha) \ge 0, \end{cases}$$

where
$$\delta(\alpha) = (\alpha^2 - 4\alpha + 4)c_2^2 - (4\alpha^2 - 11\alpha + 8)c_1c_2 + (3\alpha^2 - 7\alpha + 4)c_1^2.$$

If $\alpha \ge 2$, then
$$\eta_{RRV} = \eta_{MR}.$$

Proof. (1) The proof is similar to that of theorem 1 except that the constraint set S_1' is changed. Now we require $w_{m+1} \le \frac{1}{\alpha}b$, which implies $w^*_{m+1} = \frac{1}{\alpha}b$, and we require $W_m > \frac{\alpha-1}{\alpha}b$, which implies $r_{m+1} > c_2 - \frac{1}{\alpha}(c_2-c_1)$. Thus: $r^*_{m+1} = \sqrt{c_1c_2}$ if $\alpha \le 1+\sqrt{\frac{c_1}{c_2}}$; $r^*_{m+1} = c_2 - \frac{1}{\alpha}(c_2-c_1)$ if $\alpha > 1+\sqrt{\frac{c_1}{c_2}}$.

(2) The proof is similar to that of theorem 2 except that now we require $\frac{\alpha-1}{\alpha}b > W_m > 0$, which implies $r_{m+1} > (\alpha-1)R_m - (\alpha-2)c_1$ and thus $r^*_{m+1} = (\alpha-1)R_m - (\alpha-2)c_1$.

(3) Now, we require $\frac{1}{\alpha}b > w_{m+1} > 0$, which implies

$$R_m < c_1 + \frac{(c_2-c_1)(r_{m+1}-c_1)}{r_{m+1}-c_2+\alpha(c_2-c_1)}.$$

For $\alpha \ge 2$, the constraints $c_1 < r_{m+1} \le R_m$ make the feasible region empty, which means that V1 is dominated in the sense of worst case performance and thus can be dropped in the derivation of the worst case performance ratio, i.e., $\eta_{RRV} = \eta_{MR}$.

In the the case of $\alpha < 2$, the proof is similar to that of theorem 3 except that we now end up with

$$f_1(r_{m+1}) = \frac{(\alpha-1)c_1(c_2-c_1) + r_{m+1}(c_2-c_1)}{(\alpha-1)c_1(c_2-c_1) + (-r_{m+1}^2 + 2c_2r_{m+1} - c_1c_2)},$$

which is convex in r_{m+1} and achieves its minimum at

$r_{m+1} = -(\alpha-1)c_1 + \sqrt{\alpha(\alpha-1)c_1^2 + \alpha c_1 c_2} \equiv r'$. Since $R_m \geq r_{m+1}$, we have $r_{m+1} \leq (2-\alpha)c_2 + (\alpha-1)c_1 \equiv r''$. We note that $r'' < r'$ if and only if $\delta(\alpha) < 0$, where

$$\delta(\alpha) = (\alpha^2 - 4\alpha + 4)c_2^2 - (4\alpha^2 - 11\alpha + 8)c_1 c_2 + (3\alpha^2 - 7\alpha + 4)c_1^2.$$

Therefore,

$$\eta_{RRV} = \begin{cases} \frac{(c_2-c_1)c_2-(\alpha-1)(c_2-c_1)^2}{c_2^2-(\alpha-1)^2(c_2-c_1)^2-c_1c_2}, & \text{if } \delta(\alpha) < 0; \\ \frac{(c_2-c_1)\sqrt{\alpha(\alpha-1)c_1^2+\alpha c_1 c_2}}{2(c_2+(\alpha-1)c_2)\sqrt{\alpha(\alpha-1)c_1^2+\alpha c_1 c_2}-2\alpha c_1 c_2-(\alpha-1)(2\alpha-1)c_1^2}, & \text{if } \delta(\alpha) \geq 0. \quad \diamond \end{cases}$$

For the compound heuristics MR and RV11, it is straightforward to verify that: in the sense of worst case performance, MR dominates RV11 when $\alpha > 2$; if $1 + \sqrt{\frac{c_1}{c_2}} \leq \alpha \leq 2$, MR is better than RV11 when $\frac{c_1}{c_2} > (2 - \frac{1}{\alpha} - \frac{\alpha}{2}) - \sqrt{(2 - \frac{1}{\alpha} - \frac{\alpha}{2})^2 - (3 - \frac{2}{\alpha} - \alpha)}$, while RV11 is better than MR, otherwise; if $1 + \sqrt{\frac{c_1}{c_2}} > \alpha$, MR is better than RV11 when $\frac{c_1}{c_2} > 3 - \sqrt{2}$, while RV11 is better than MR when $\frac{c_1}{c_2} < 3 - \sqrt{2}$.

Thus far, we have assumed that $w_j \leq b$, $\forall\, j$. If we relax this assumption, then α can be less than 1 (but we assume $\alpha > 0$). In that case, η_{R1} and η_{V1} are the same as the case of $\alpha = 1$. Since, in R1 and V1, we always underuse the prespecified capacity, any items exceeding the prespecified capacity will not be chosen. In the worst case, we can only choose an arbitrarily small item: then η_{R1} and η_{V1} approach $\frac{c_1}{c_2}$ in the limit since we can sell the unused capacity for up to bc_1, while the maximum possible profit we could make is bc_2. For R2, the results of lemma 5 apply since, in its proof, the requirement is $\alpha > 0$. For V2, η_{V2} is the same as the case of $\alpha = 1$ if $\frac{1}{2} < \alpha < 1$; $\eta_{V2} = \max\{\frac{c_1}{\alpha c_2} - (\frac{1}{\alpha} - 1),\ 0\}$ if $0 < \alpha < \frac{1}{2}$. For MR and RV11, the results of theorem 4 apply since η_{MR} and η_{RV11} are independent of α when $\alpha \leq 1 + \sqrt{\frac{c_1}{c_2}}$ and $\alpha < 2$, respectively. For RRV, the results of theorem 4 apply since $1 > \alpha > 0$ causes no change in the proof.

5. A Note on Other Flexible Knapsack Problems

For FSSP, we assume that $w_1 \geq w_2 \geq \ldots \geq w_n$ and consider the following two greedy algorithms, W1 and W2.

W1: 1. Set $j = 1$ and $B = b$.
2. If $w_j \leq B$ then set $x_j = 1$ and $B = B - w_j$; otherwise, set $x_j = 0$.
3. Stop if $j = n$ or $B = 0$; otherwise, set $j = j + 1$ and go to step 2.

W2: 1. Set $j = 1$ and $B = b$.
2. If $w_j \leq B$ then set $x_j = 1$ and $B = B - w_j$; otherwise, set $x_j = 1$ and stop.
3. Set $j = j + 1$ and go to step 2.

Let MR denote the compound heuristic of W1 and W2. Define $\alpha = \min_i\{\frac{b}{w_i}\}$. We can adopt the notation defined for R1 and R2 in FKP to derive η_{MW}. Here, we have $Z^{W1} \geq (1 - c_1)W_m + bc_1$, $Z^{W2} \geq bc_2 - (W_m + w_{m+1})(c_2 - 1)$, $Z^* \geq b$, $R_m = r_{m+1} = 1$, and the constraint set $S_{MR} = \{w_m \geq w_{m+1},\ w_{m+1} \leq \frac{1}{\alpha}b,\ W_m > \frac{\alpha-1}{\alpha}b\}$. Hence, similar to the proof in theorem 1, we have $w^*_{m+1} = W^*_m = \frac{c_2 - c_1}{c_2 - c_1 + c_2 - 1} b$ if $\alpha \le 2$, and $w^*_{m+1} = \frac{b}{\alpha}$, $W^*_m = (1 - \frac{c_2 - 1}{\alpha(c_2 - c_1)}) b$ if $\alpha > 2$; and thus

$$\eta_{MW} = \begin{cases} \frac{(c_2 - c_1) + c_1(c_2 - 1)}{(c_2 - c_1) + (c_2 - 1)}, & \text{if } \alpha \le 2; \\ 1 - \frac{1}{\alpha} \frac{(c_2 - 1)(1 - c_1)}{c_2 - c_1}, & \text{if } \alpha > 2. \end{cases}$$

Note that if $c_2 = 1.1$, $c_1 = .9$, we have $\eta_{MW} = \frac{29}{30}$ if $\alpha \le 2$ and $\eta_{MW} = \frac{59}{60}$ if $\alpha = 3$.

For FBKP, we assume that $r_1 \ge r_2 \ge \ldots \ge r_n$ and consider the following two greedy algorithms, R1 and R2.

R1: 1. Set $j = 1$ and $B = b$.
  2. Set $x_j = \lfloor \frac{B}{w_j} \rfloor$ and $B = B - x_j w_j$.
  3. Stop if $j = n$ or $B = 0$; otherwise, set $j := j + 1$ and go to step 2.

R2: 1. Set $j = 1$ and $B = b$.
  2. Set $x_j = \lfloor \frac{B}{w_j} \rfloor$ and $B = B - x_j w_j$.
  If $x_j \ge 1$, go to setp 3; otherwise, set $x_j = 1$ and stop.
  3. Set $j = j + 1$ and go to step 2.

Define $f_i = b \pmod{w_i}$, $\alpha_i = \frac{b}{w_i} - f_i$, and $\alpha' = \min\{\alpha_1, u_1\}$. We can adopt the notation defined for R1 and R2 in KP if we interpret item m as the m-th unit item instead of the m-th item. Here, we have Z^{LP}, Z^{R1} and Z^{R2} as defined in KP, and the constraint set $S_{MR} = \{c_2 > r_{m+1} > c_1,\ b \ge W_m > 0,\ W_m > \frac{\alpha'}{\alpha_1 + f_1} b,\ w_{m+1} \le \frac{\alpha_1 + f_1 - \alpha'_1}{\alpha_1 + f_1}\}$. Hence, similar to the proof in theorem 1, we have $R^*_m = r^*_{m+1} = c_2$, $W^*_m = \frac{\alpha'}{\alpha_1 + f_1} b$, $w^*_{m+1} = \frac{\alpha_1 + f_1 - \alpha'}{\alpha_1 + f_1} b$ and thus

$$\eta_{MR} = \frac{c_1}{c_2} + \frac{\alpha'}{\alpha_1 + f_1}.$$

For FUKP, we have $\eta_{MR} = \frac{c_1}{c_2} + \frac{\alpha_1}{\alpha_1 + f_1}$, which is obtained by simply setting $\alpha' = \alpha_1$.

6. Conclusions

Lemma 3 in section 3 is useful for analyzing a compound heuristic in general. One strength of this optimization approach is that its derivation of a worst case performance ratio is suggestive in the construction of the worst case examples. For example, in theorems 1, 2 and 3 (in section 3) once we know R^*_m, r^*_{m+1}, W^*_m, w^*_{m+1}, the construction of a worst case example is straightforward. Some steps in this optimization approach remain to be a matter of design, e.g., the choice of surrogate algorithms and the choice of surrogate admissible regions. To provide some guidelines regarding these choices, more successful examples of using this optimization approach to analyze other heuristics are needed.

References

[1] Fisher, M. L., "Worst Case Analysis of Heuristic Algorithms" *Management Science* **26**, 1–17 (1980).

[2] Frederickson, G., Hecht, M., and Kim, C., "Approximation Algorithms for Some Routing Problems," *SIAM J. Computing* **7**, 178–193 (1978).

[3] Garey, M. R., and Johnson, D. S., *Computers and Intractability: A Guide to the Theory of NP-completeness*. Freeman, San Francisco, California, 1979.

[4] Langston, M. A., "A Study of Composite Heuristic Algorithms" *J. of Operational*

Research Society. **38**, 539–544 (1987).

[5] Lueker, G. S., "Two NP-complete Problems in Nonnegative Integer Programming," Report No. 178, Computer Science Laboratory, Princeton University, Princeton, NJ, 1975.

[6] Martello, S., and Toth, P., *Knapsack Problems: Algorithms and Computer Implementations*, Wiley, Chichester, 1990.

[7] Yao, A. C., "New Algorithms for Bin Packing," *J. Ass. Comput. Mach.* **27**, 207-227, (1980).

Tsung-Chyan LAI: College of Management,
National Taiwan University, 50 Lane 144,
Keelung Road Section 4, Taipei 106, TAIWAN