Strengthening Landmark Heuristics via Hitting Sets
Technical Report: Proofs

Blai Bonet
Departamento de Computación
Universidad Simón Bolívar
Caracas, Venezuela
bonet@ldc.usb.ve

Malte Helmert
Albert-Ludwigs-Universität Freiburg
Institut für Informatik
Georges-Köhler-Allee 52
79110 Freiburg, Germany
helmert@informatik.uni-freiburg.de

Hitting Sets

Let $A = \{a_1, \ldots, a_n\}$ be a set and $\mathcal{F} = \{F_1, \ldots, F_m\}$ a family of subsets of A. A subset $H \subseteq A$ has the hitting set property, or is a hitting set, iff $H \cap F_i \neq \emptyset$ for $1 \leq i \leq m$ (i.e., H "hits" each set F_i). If we are given a cost function $c : A \to \mathbb{N}$, the cost of H is $\sum_{a \in H} c(a)$. A hitting set is of minimum cost if its cost is minimal among all hitting sets.

The problem of finding a minimum-cost hitting set for family $\mathcal{F}$ and cost function c is denoted by $\langle \mathcal{F}, c \rangle$, and the cost of its solution by $\min(\mathcal{F}, c)$. A relaxation for $\langle \mathcal{F}, c \rangle$ is a problem $\langle \mathcal{F}', c' \rangle$ such that $c' \leq c$, and for all $F' \in \mathcal{F}'$ there is $F \in \mathcal{F}$ with $F \subseteq F'$. In words, $\langle \mathcal{F}, c \rangle$ can be relaxed by reducing costs, dropping sets from $\mathcal{F}$, or enlarging elements of $\mathcal{F}$. Determining the existence of a hitting set for a given cost bound is a classic problem in computer science, one of the first problems to be shown NP-complete (Kar72).

Lemma 1. *If $\langle \mathcal{F}', c' \rangle$ is a relaxation of $\langle \mathcal{F}, c \rangle$, then $\min(\mathcal{F}', c') \leq \min(\mathcal{F}, c)$. Furthermore, if $\{\langle \mathcal{F}_i, c_i \rangle\}$ is a collection of relaxations of $\mathcal{F}$ such that $\sum_i c_i \leq c$, then $\sum_i \min(\mathcal{F}_i, c_i) \leq \min(\mathcal{F}, c)$.*

Proof. The first claim is direct since a hitting set for $\mathcal{F}$ is also a hitting set for $\mathcal{F}'$ and $c' \leq c$. For the second claim, consider a hitting set H for $\mathcal{A}$. Then,

$$c(H) \geq \sum_i c_i(H) \geq \sum_i \min(\mathcal{F}_i, c_i) \,.$$

The first inequality holds because $\sum_i c_i \leq c$ and the second because H is a hitting set for each $\mathcal{F}_i$. □

Decomposition and Width

Let $\mathcal{F}$ be a family that can be partitioned into $\Pi = \{\mathcal{F}_1, \ldots, \mathcal{F}_m\}$ satisfying $(\bigcup \mathcal{F}_i) \cap (\bigcup \mathcal{F}_j) = \emptyset$ for all $i \neq j$; i.e., the blocks in the partition are pairwise *independent*. Then, for any cost function c, $\min(\mathcal{F}, c) = \sum_{i=1}^{m} \min(\mathcal{F}_i, c)$ and the problem of finding a minimum-cost hitting set for $\mathcal{F}$ can be decomposed into smaller subproblems. We call the maximum size of a block in Π the *width* of Π, denoted by *width*(Π). The *width* of $\mathcal{F}$, denoted by *width*$(\mathcal{F})$ is the minimum *width*(Π) over all partitions Π of $\mathcal{F}$ into independent blocks. Finding a partition that minimizes the width is an easy problem similar to computing connected components of a graph, so *width*$(\mathcal{F})$ can be efficiently computed.

Indeed, for a family $\mathcal{F}$, define the (undirected) graph $G = \langle V, E \rangle$ where $G = \bigcup \mathcal{F}$ and $\{a, a'\} \in E$ iff $a \neq a'$ and there is $F \in \mathcal{F}$ with $F \supseteq \{a, a'\}$. Then,

Proposition 1. *The width of $\mathcal{F}$ equals the size of the largest connected component of G.*

Proof. Let $\{G_1, \ldots, G_m\}$ be the connected components of G and w the size of a largest one. We need to show that *width*$(\mathcal{F}) \leq w$ and $w \leq$ *width*$(\mathcal{F})$.

For the first inequality, consider the partition $\Pi_G = \{\mathcal{F}_1, \ldots, \mathcal{F}_m\}$ of $\mathcal{F}$ defined by $\mathcal{F}_i = \{F \in \mathcal{F} : F \cap G_i \neq \emptyset\}$. It is easy to check that this partition is well defined and with pairwise independent blocks. Therefore,

$$\textit{width}(\mathcal{F}) \leq \textit{width}(\Pi_G) = \max_{i,\ldots,m} |\mathcal{F}_i| = w \,.$$

For the second inequality, let $\Pi = \{\mathcal{F}_1, \ldots, \mathcal{F}_k\}$ be a partition of $\mathcal{F}$ into pairwise independent blocks. It will be enough to show that each connected component G_i is contained in some block of Π.

Let a be a vertex in G_i and $\mathcal{F}_j \in \Pi$ a block containing it; i.e., $a \in \bigcup \mathcal{F}_j$. Consider an arbitrary vertex $a' \in G_i$. Since G_i is connected, there is a path $\langle a = a_0, a_1, \ldots, a_\ell = a' \rangle$ in G_i. We perform induction on ℓ to show that a' is also in $\bigcup \mathcal{F}_j$. For $\ell = 0$, $a = a'$ and the claim is trivial. Assume that the claim holds for paths of length less than ℓ. Then, $a_{\ell-1}$ is in $\bigcup \mathcal{F}_j$ by inductive hypothesis. This implies that there is an $F \in \mathcal{F}_j$ such that $a_{\ell-1} \in F$. On the other hand, the edge $\{a_{\ell-1}, a_\ell\}$ implies the existence of F' with $\{a_{\ell-1}, a_\ell\} \subseteq F'$. Therefore, $F \cap F' \neq \emptyset$, $F' \in \mathcal{F}_j$ (because the blocks of Π are pairwise independent), and $a_\ell \in \bigcup \mathcal{F}_j$. □

Let $\mathcal{F} = \{F_1, \ldots, F_k\}$ be a family over A with k subsets, but with no assumptions on the sizes of each F_i or A. We show that $\min(\mathcal{F}, c)$ and a hitting set achieving this cost can be computed in time bounded by $O(\|\mathcal{F}\| + k4^k)$. To see this, consider the hypergraph $H_{\mathcal{F}} = (X, E)$ where $X = \{1, \ldots, k\}$ and there is a hyperedge $e(a) = \{i : a \in F_i\}$ with cost $c(a)$ for each $a \in A$. The hitting sets for $\mathcal{F}$ are in one-to-one correspondence with the covers of the hypergraph $H_{\mathcal{F}}$ (a cover is a set of hyperedges that "touch" every vertex). Hence, finding $\min(\mathcal{F}, c)$ is equivalent to finding a minimum-cost cover for $H_{\mathcal{F}}$. For the latter, observe that all hyperedges $e(a)$ for which there is a hyperedge $e(a')$ with

Input: Hypergraph $H = (X, E, c)$ where $X = \{1, \ldots, k\}$, c is the edge-cost function, and $|E| \leq 2^k$.
Output: $\texttt{best\_cover}[X']$, for $X' \subseteq X$, is the cost of a minimum-cost cover for the vertices in X'.
Thus, $\texttt{best\_cover}[X]$ is the cost of a minimum-cost cover for H.

$\texttt{best\_cover}[\emptyset] := 0$
forall $X' \subseteq \{1, \ldots, k\}$ with $X' \neq \emptyset$ **do**
    $\texttt{best\_cover}[X'] := \infty$

forall $X' \subseteq \{1, \ldots, k\}$ in order of increasing cardinality **do**
    forall hyperedges $e \in E$ of the graph **do**
        $\texttt{best\_cover}[X'] := \min(\texttt{best\_cover}[X'], c(e) + \texttt{best\_cover}[X' \setminus \{e\}])$

Figure 1: DP algorithm for computing a minimum-cost cover of an hypergraph.

$e(a) = e(a')$ and $c(a') < c(a)$ may be removed (and if $c(a') = c(a)$, only one of the hyperedges needs to be kept).

Since a hyperedge is a subset of X, this implies that we only need to consider hypergraphs with at most 2^k edges. Using dynamic programming, a minimum cost cover for such a hypergraph can be found in time $O(k4^k)$. Combining this with the time required for constructing the hypergraph yields the overall $O(\|\mathcal{F}\| + k4^k)$ bound. This is an example of *fixed-parameter tractability* (FG06).

Theorem 2. *The problem of computing* $\min(\mathcal{F}, c)$ *is fixed-parameter tractable when considering the width of* $\mathcal{F}$ *as the parameter. In particular, for any* fixed *bound* k*,* $\min(\mathcal{F}, c)$ *for families of width at most* k *can be computed in linear time.*

Proof. We first show that a hitting set problem $\langle \mathcal{F} = \{F_1, \ldots, F_k\}, c\rangle$, of arbitrary width, can be solved in time $O(\|\mathcal{F}\| + k4^k)$, where $\|\mathcal{F}\|$ denotes the input size for F.

Consider the hypergraph $H_{\mathcal{F}} = (X, E)$ where $X = \{1, \ldots, k\}$ and there is a hyperedge $e(a) = \{i : a \in F_i\}$ with cost $c(a)$ for each $a \in A$. The hitting sets for $\mathcal{F}$ are in one-to-one correspondence with the covers of $H_{\mathcal{F}}$ (a cover is a set of hyperedges that "touch" every vertex). Hence, finding $\min(\mathcal{F}, c)$ is equivalent to finding a minimum-cost cover for $H_{\mathcal{F}}$. For the latter, observe that all hyperedges $e(a)$ for which there is a hyperedge $e(a')$ with $e(a) = e(a')$ and $c(a') < c(a)$ may be removed (if $c(a') = c(a)$, only one of the hyperedges need to be kept). Thus, we only need to consider hypergraphs with at most 2^k edges. The following DP algorithm computes the cost of a minimum-cost cover.

We make use of a table $\texttt{best\_cover}$ of size 2^k that maps subsets of $X = \{1, \ldots, k\}$ into $[0, \infty) \cup \{\infty\}$. At the end of the algorithm, the entry $\texttt{best\_cover}[X]$ contains the cost of a minimum-cost cover for $\langle \mathcal{F}, c\rangle$, and a minimum-cost cover can be recovered from the table in linear time. The algorithm initializes the table as $\texttt{best\_cover}[\emptyset] := 0$ and $\texttt{best\_cover}[X'] := \infty$ for all $\emptyset \subset X' \subseteq X$. Then, it updates the table using DP as shown in Fig. 1. The initialization loop takes $O(2^k)$ time. The two nested loops make a total of $O(4^k)$ iterations, each taking time $O(k)$. The running time of the DP algorithm is $O(\|\mathcal{F}\|)$ time for constructing the hypergraph, and $O(k4^k)$ time for finding a minimum-cost cover of the hypergraph.

Now, we show that a hitting set problem $\langle \mathcal{F}, c\rangle$ of width at most k can be solved in linear time. By Proposition 1, one can compute in linear time a partition $\Pi = \{\mathcal{F}_1, \ldots, \mathcal{F}_m\}$ of $\mathcal{F}$ into independent blocks. Then, since the blocks are independent, $\min(\mathcal{F}, c) = \sum_{i=1}^{m} \min(\mathcal{F}_i, c)$. Each subproblem $\langle \mathcal{F}_i, c\rangle$ can be solved in time $O(\|\mathcal{F}_i\| + k4^k)$. Therefore, $\langle \mathcal{F}, c\rangle$ can be solved in time

$$\begin{aligned}\sum_{i=1}^{m} O(\|\mathcal{F}_i\| + k4^k) &= O(m\|\mathcal{F}_i\| + mk4^k) \\ &= O(\|\mathcal{F}\| + \|\mathcal{F}\|k4^k)\end{aligned}$$

which is linear in $\|\mathcal{F}\|$ since k is fixed and bounded. □

Landmarks

A STRIPS problem with action costs is a tuple $P = \langle F, O, I, G, c\rangle$ where F is the set of fluents, O is the set of actions or operators, I and G are the initial state and goal description, and $c : O \to \mathbb{N}$ is the cost function. We are interested in delete relaxations, so we assume that the operators have empty delete lists, and thus 'plan' and 'relaxed plan' shall denote the same. For a definition of the basic concepts underlying delete relaxations, such as the $h^{\max}$ and h^+ functions, we refer to the literature (HD09). We also assume from now on that all fluents have finite $h^{\max}$ values, which implies that the problem has finite h^+ value. As additional simplifying assumptions, we require that all operators have nonempty preconditions, that there are two fluents $s, t \in F$ such that $I = \{s\}$ and $G = \{t\}$, and that there is a unique operator *fin* that adds t. When these simplifying assumptions are not met, they can be achieved through simple linear-time transformations. We denote the precondition and effects of $a \in O$ by pre(a) and post(a). The h^+ value for state I is denoted by $h^+(P)$.

An (action) landmark for P is a disjunction $a_1 \vee \cdots \vee a_n$ of actions such that every plan for P must contain at least one such action. Such a landmark is denoted by the set $\{a_1, \ldots, a_n\}$.

Recall that a pcf D assigns a precondition $D(a) \in \text{pre}(a)$ to each action $a \in O$. Our first result relates cuts in the justification graph $G(D)$ with landmarks for P.

Lemma 3. *Let D be a pcf and C an s-t-cut of $G(D)$. Then, the labels of the edges in the cut-set of C form a landmark.*

Proof. A relaxed plan defines an s-t-path on $G(D)$ that must cross every s-t-cut. $\square$

Given a pcf D, we denote the set of landmarks associated with the cut-sets of $G(D)$ by Landmarks(D). By considering all pcfs and all cuts in the justification graphs, we obtain the hitting set problem $\mathcal{F}_L \doteq \bigcup\{\text{Landmarks}(D) : D \text{ is a precondition-choice function}\}$.

Theorem 4. *If H is a plan for P, then H is a hitting set for $\mathcal{F}_L$. Conversely, if H is a hitting set for $\mathcal{F}_L$, then H contains a plan for P. Therefore, $\min(\mathcal{F}_L, c) = h^+(P)$.*

Proof. The first claim is direct, since by Lemma 3, every element of $\mathcal{F}_L$ is hit by every plan. The last claim follows from the first two.

For the second claim, let H be a hitting set for $\mathcal{F}_L$ and let R be the set of fluents that can be reached by only using operators in H. If R contains the goal t, then H contains a plan and there is nothing to prove. So, assume $t \notin R$. We construct a pcf D such that $G(D)$ contains an s-t-cut whose cut-set is not hit by H, thus reaching a contradiction. We classify operators into three types and define D:

- T1. If $\text{pre}(a) \subseteq R$ and $\text{post}(a) \subseteq R$, then set $D(a)$ arbitrarily to some $p \in \text{pre}(a)$.
- T2. If $\text{pre}(a) \subseteq R$ and $\text{post}(a) \nsubseteq R$, then set $D(a)$ arbitrarily to some $p \in \text{pre}(a)$.
- T3. If $\text{pre}(a) \nsubseteq R$, then set $D(a)$ to some $p \in \text{pre}(a) \setminus R$.

Now consider the cut (R, R^c) of $G(D)$, where R^c is the set of all fluents not in R. It is a cut since $s \in R$ and $t \notin R$. We show that H does not hit the cut-set, i.e., there exists no operator $a \in H$ that labels an edge going from some fluent in R to some fluent not in R.

Assume that $a \in H$ were such an operator. It cannot be of type T1, because edges labeled by type T1 operators go from R into R. It cannot be of type T2, because $\text{pre}(a) \subseteq R$ and $a \in H$ implies $\text{post}(a) \subseteq R$ (by definition of R). Finally, it cannot be of type T3, as edges labeled by type T3 operators do not start in R. Hence, no such operator exists. $\square$

In practice, computing $\mathcal{F}_L$ according to the definition above is infeasible because there are usually exponentially many pcfs. However, if we can compute and solve, in polynomial time, a relaxation of $\mathcal{F}_L$, then this provides a polytime admissible approximation of h^+.

Corollary 5. *Let $\langle \mathcal{F}, c' \rangle$ be a polynomial-time computable relaxation of $\langle \mathcal{F}_L, c \rangle$ (possibly additive[1]) whose solution is polynomial-time computable. Then the heuristic $h = \min(\mathcal{F}, c')$ is a polytime admissible approximation of h^+.*

[1] In the additive case, we slightly abuse notation since $\langle \mathcal{F}, c' \rangle$ should be replaced by a collection $\{\langle \mathcal{F}_i, c_i \rangle\}_i$.

Proof. Direct from the assumptions: the admissibility because $\langle \mathcal{F}, c' \rangle$ is a relaxation of $\langle \mathcal{F}_L, c \rangle$, which defines h^+, and the polynomial-time computability because $\langle \mathcal{F}, c' \rangle$ is solvable in polynomial time. $\square$

An important special case covered by the corollary are landmark heuristics based on cost partitioning, including LM-cut (see below) and the heuristics of Karpas and Domshlak (KD09). In general, given a set $\mathcal{L} = \{L_1, \ldots, L_n\}$ of landmarks, a cost partitioning for $\mathcal{L}$ is a collection $\mathcal{C} = \{c_1, \ldots, c_n\}$ of cost functions such that $\sum_{i=1}^{n} c_i(a) \leq c(a)$ for each action a. The partitioning defines the heuristic $h_{\mathcal{C}} \doteq \sum_{i=1}^{n} \min_{a \in L_i} c_i(a)$, which is an additive relaxation of $\mathcal{F}_L$ when $\mathcal{L} \subseteq \mathcal{F}_L$.

Karpas and Domshlak studied *uniform* cost partitioning, defined as $c_i(a) \doteq 0$ if $a \notin L_i$ and $c_i(a) \doteq c(a)/|\{i : a \in L_i\}|$ if $a \in L_i$, and *optimal* cost partitioning, which maximizes $h_{\mathcal{C}}$ through linear programming (LP). Interestingly, there is a close connection between the optimal cost partitioning LP and the hitting set ILP for $\mathcal{L}$.

Theorem 6. *Let $\mathcal{L}$ be a collection of landmarks, and let c be the cost function for the actions. Then, the LP that defines the optimal cost partitioning is the* dual *of the LP relaxation of the ILP for $\langle \mathcal{L}, c \rangle$.*

Proof. Let $\mathcal{L} = \{L_j\}_j$ be a collection of landmarks over actions A, and c a cost functions for the actions. The ILP corresponding to $\langle L, c \rangle$ is:

$$\begin{aligned} &\text{minimize} \quad \textstyle\sum_{a \in A} c(a)\, x_a \\ &\text{subject to} \\ &(\text{for } L_j \in \mathcal{L}) \quad \textstyle\sum_{a \in L_j} x_a \geq 1\,, \\ &(\text{for } a \in A) \quad x_a \in \{0, 1\}\,. \end{aligned}$$

The variables x_a are indicator variables that define the hitting set. The LP relaxation of the ILP is:

$$\begin{aligned} &\text{minimize} \quad \textstyle\sum_{a \in A} c(a)\, x_a \\ &\text{subject to} \\ &(\text{for } L_j \in \mathcal{L}) \quad \textstyle\sum_{a \in L_j} x_a \geq 1\,, \\ &(\text{for } a \in A) \quad 0 \leq x_a \leq 1\,. \end{aligned}$$

The dual of the LP is:

$$\begin{aligned} &\text{maximize} \quad \textstyle\sum_{j} y_j \\ &\text{subject to} \\ &(\text{for } a \in A) \quad \textstyle\sum_{j} [\![a \in L_j]\!] y_j \leq c(a)\,, \\ &(\text{for } L_j \in \mathcal{L}) \quad y_j \geq 0\,. \end{aligned}$$

The variables y_j are the dual variables corresponding to the constraints of the LP. This LP attains the same value of the LP that defines the optimal cost partitioning for $\mathcal{L}$. Indeed, the y_j variables can be interpreted as the cost of each landmark for the optimal cost assignment. The constraints avoid cost partitionings c_j such that $\sum_j c_j(a) \geq c(a)$. $\square$

The LM-Cut Heuristic

Theorem 7. *Given a set of landmark $\mathcal{L} \subseteq \mathcal{F}_L$ and cost partitioning $\mathcal{C}$, $h_{\mathcal{C}}$ is an additive relaxation of $\mathcal{F}_L$. LM-cut is one such relaxation.*

Proof. The first claim is direct by definition. For the second claim, we show that there is a collection $\{\mathcal{F}_i\}_{i=1}^n$ of relaxations of $\mathcal{F}_L$ such that

$$h^{\text{LM-cut}}(P) = \sum_{i=1}^{n} m_i = \sum_{i=1}^{n} \min(\mathcal{F}_i, c_i) \leq \min(\mathcal{F}_L, c)\,.$$

Let $L_1, \ldots, L_n$ and $c_1, \ldots, c_n$ be the landmarks and cost functions computed by LM-cut at each stage. Define $\mathcal{F}_i = \{L_i\}$ and $c'_i(a) = m_i$ if $a \in L_i$ and $c'_i(a) = 0$ otherwise. Clearly, each $\mathcal{F}_i$ is a relaxation of $\mathcal{F}_L$ and $\min(\mathcal{F}_i, c_i) = \min(\mathcal{F}_i, c'_i) = m_i$. By Lemma 1, it remains to show that $\sum_i c'_i \leq c$.

Let $I(a,k) = \{i : a \in L_i, 1 \leq i \leq k\}$ be the set of indices for the landmarks in $\{L_1, \ldots, L_k\}$ that contain a. Using induction it is not difficult to show that $c_{k+1}(a) = c(a) - \sum_{i \in I(a,k)} m_i$. On the other hand,

$$c'_1(a) + c'_2(a) + \cdots + c'_k(a) = \sum_{i \in I(a,k)} c'_i(a) = \sum_{i \in I(a,k)} m_i\,.$$

We show using induction on k that $c'_1(a) + \cdots + c'_k(a) \leq c(a)$ for every action a and $1 \leq k \leq n$. The base of the induction is easy. If $a \notin L_1$, then $c'_1(a) = 0 \leq c(a)$. If $a \in L_1$, then $c'_1(a) = m_1 = \min_{a' \in L_1} c_1(a') \leq c(a)$. Suppose that the claim holds up to k. We need to show it for $k+1$. If $a \notin L_{k+1}$, then $\sum_{i \in I(a,k+1)} c'_i(a) = \sum_{i \in I(a,k)} c'_i(a) \leq c(a)$ by inductive hypothesis. If $a \in L_{k+1}$, then

$$\begin{aligned}
\sum_{i \in I(a,k+1)} c'_i(a) &= \sum_{i \in I(a,k)} c'_i(a) + c'_{k+1}(a) \\
&= \sum_{i \in I(a,k)} m_i + m_{k+1} \\
&= \sum_{i \in I(a,k)} m_i + \min_{a' \in L_{k+1}} c_{k+1}(a') \\
&\leq \sum_{i \in I(a,k)} m_i + c_{k+1}(a) \\
&= \sum_{i \in I(a,k)} m_i + c(a) - \sum_{i \in I(a,k)} m_i \\
&= c(a)\,.
\end{aligned}$$

$\square$

Theorem 8. *For any fixed $p \geq 1$ and $k \geq 1$, $h_{p,k}^{\text{LM-cut}}$ is computable in polynomial time and dominates $h^{\text{LM-cut}}$.*

Proof. Direct. $\square$

References

Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Springer-Verlag, 2006.

Malte Helmert and Carmel Domshlak. Landmarks, critical paths and abstractions: What's the difference anyway? In *Proc. ICAPS 2009*, pages 162–169, 2009.

Richard M. Karp. Reducibility among combinatorial problems. In Raymond E. Miller and James W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. 1972.

Erez Karpas and Carmel Domshlak. Cost-optimal planning with landmarks. In *Proc. IJCAI 2009*, pages 1728–1733, 2009.