Chapter S:VI

VI. Relaxed Models

- ❑ Motivation
- ❑ ε-Admissible Speedup Versions of A*
- ❑ Using Information about Uncertainty of h
- ❑ Risk Measures

- ❑ Nonadditive Evaluation Functions

- ❑ Heuristics Provided by Simplified Models
- ❑ Mechanical Generation of Admissible Heuristics
- ❑ Probability-Based Heuristics

Using Information about Uncertainty of *h*

Using a non-admissible heuristic function

Idea [Harris]:

The heuristic function h estimates the cheapest remaining cost h^* mostly quite well, but sometimes overestimates h^* by no more than ε.

➜ A* using such a heuristic function h is ε-admissible.

The condition for ε-admissibility of A* is satisfied because at termination it holds

$$h(n) - h^*(n) \leq \varepsilon \text{ für alle } n \in \text{ OPEN.}$$

Also the weakened form of admissibility of h is often too restrictive.

Often it is easier to find a heuristic estimate for h^* that mostly estimates precisely but sometimes overestimates h^* (by much more than any reasonable ε).

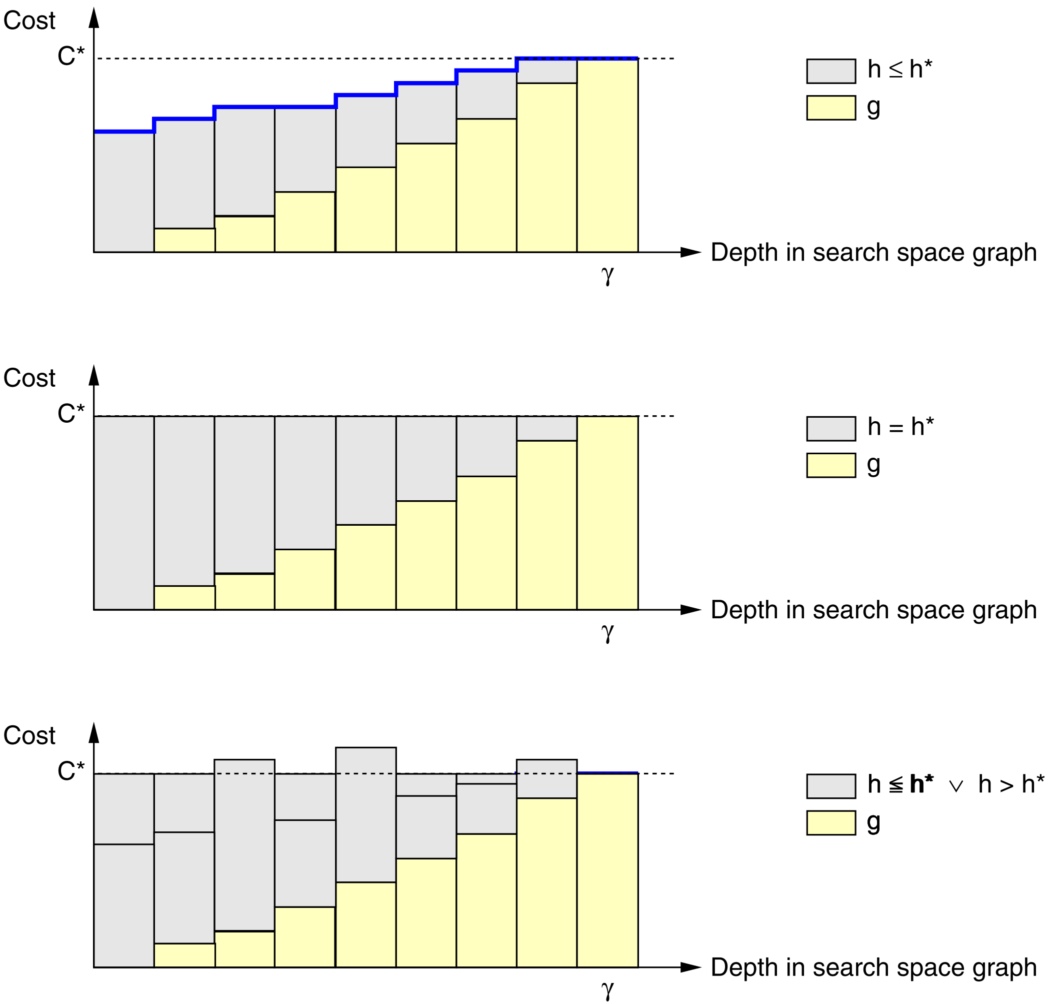
➜ The error in the estimate is not limited, but a large error is unlikely.

Remarks:

- Heuristic functions h with $h \leq (1+\varepsilon)h^*$ are called ε-admissible.
- Analogously to Lemma C^*-Bounded OPEN Node, it can be proven that, at any point in time before termination, there exists some node n in OPEN with $f(n) \leq (1+\varepsilon)C^*$.
- The condition "$h(n) - h^*(n) \leq \varepsilon$ for all $n \in$ OPEN" is sufficient, but not necessary, for A* being ε-admissible.

Using Information about Uncertainty of *h*

Illustration of Underestimating and Overestimating Estimation Functions



Using Information about Uncertainty of *h*

Example: Search in "Random" Graphs

Given is a graph with randomly drawn edge costs. The minimum number of edges to a target node is known in each node.

- Edge costs $c(n, n')$ are known to be drawn independently from a common distribution function, uniform in interval $[0; 1]$.
- For long paths with N edges from a node n to a goal node in Γ it is known that h*(n) is most likely to be near $\frac{N}{2}$.
- The only *admissible* heuristic estimate for h^* is $h_1(n) = 0$.
- The most reasonable heuristic estimate for h^* is $h_2(n) = \frac{N}{2}$.

The heuristic estimate h_2 leads to a worst-case cost overestimation of $\frac{N}{2}$ and is therefore not (ε-)admissible. But the likelihood of this event is extremely small.

Using Information about Uncertainty of *h*

Example: Search in “Random” Graphs

Given is a graph with randomly drawn edge costs. The minimum number of edges to a target node is known in each node.

- Edge costs $c(n, n')$ are known to be drawn independently from a common distribution function, uniform in interval $[0; 1]$.
- For long paths with N edges from a node n to a goal node in Γ it is known that h*(n) is most likely to be near $\frac{N}{2}$.
- The only *admissible* heuristic estimate for h^* is $h_1(n) = 0$.
- The most reasonable heuristic estimate for h^* is $h_2(n) = \frac{N}{2}$.

The heuristic estimate h_2 leads to a worst-case cost overestimation of $\frac{N}{2}$ and is therefore not (ε-)admissible. But the likelihood of this event is extremely small.

→ Algorithm R*$_\delta$:

– Besides an estimation function h for h^* there is also knowledge about the uncertainty of the estimation process.

– Knowledge about the uncertainty of the estimation process is expressed in the form of a probability density function $\rho_{h^*}(x)$.

Using Information about Uncertainty of *h*

Describing the Estimation Uncertainty using Density Functions

Viewing cost functions as a random variables:

cost function	**random variable**
$h^*(n)$	h_n^*
$f^*(n) = g^*(n) + h^*(n)$	f_n^*
$f^+(n) = g(n) + h^*(n)$	f_n^+

Let $\rho_{h_n^*}$ be a density function for the random variable h_n^*.

Semantics:

On the basis of $\rho_{h_n^*}$ one can define the probability with which $h^*(n)$ can be found in a neighborhood of x costs.

$$P(h_n^* = x) = \rho_{h_n^*}(x)$$

Using Information about Uncertainty of *h*

Describing the Estimation Uncertainty using Density Functions (continued)

Let $\rho_{h_n^*}$ be a density function for the random variable h_n^*.

Further applies:

1. From $\rho_{h_n^*}(x)$ a density function $\rho_{f_n^*}(y)$ can be derived for the random variable f_n^*, if g^* is known (e.g., when searching a tree):

   $$\rho_{f_n^*}(y) := \rho_{h_n^*}(y - g^*)$$

2. Let P_{s-n} be the cheapest known path from s to an OPEN node n.
   From $\rho_{h_n^*}(x)$ a density function $\rho_{f_n^+}(y)$ can be derived for the random variable f_n^+, which specifies the cost of an optimal solution path that continues P_{s-n}:

   $$\rho_{f_n^+}(y) := \rho_{h_n^*}(y - g)$$

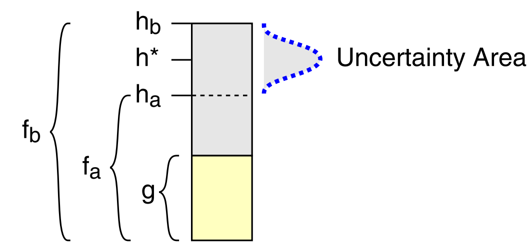
Remarks:

- The random variable f_n^+ with associated density function $\rho_{f_n^+}$ is given for each node n.
- The random variable f_n^+ describes the possible costs of an optimal solution path that contains the pointer path P_{s-n} as a subpath.
- If goal nodes can be reached from s, the OPEN list always contains a node n, to which $f^+(n) = f^*(n)$ applies. [Corollrary Shallowest OPEN Node on Optimum Path]

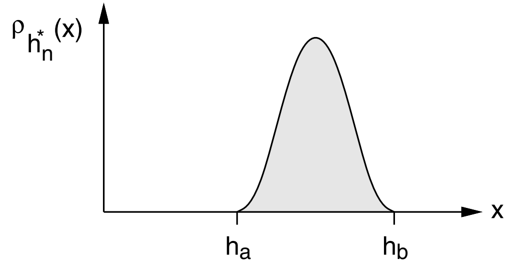
Using Information about Uncertainty of *h*

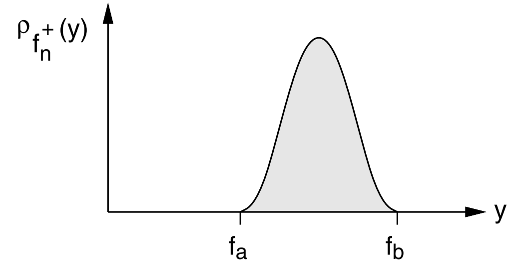
Describing the Estimation Uncertainty using Density Functions (continued)

Uncertainty area:

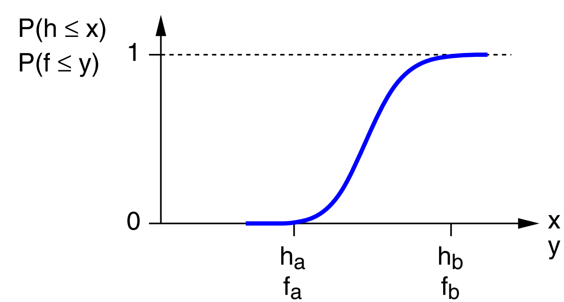


Density functions ρ:





Related distribution function:

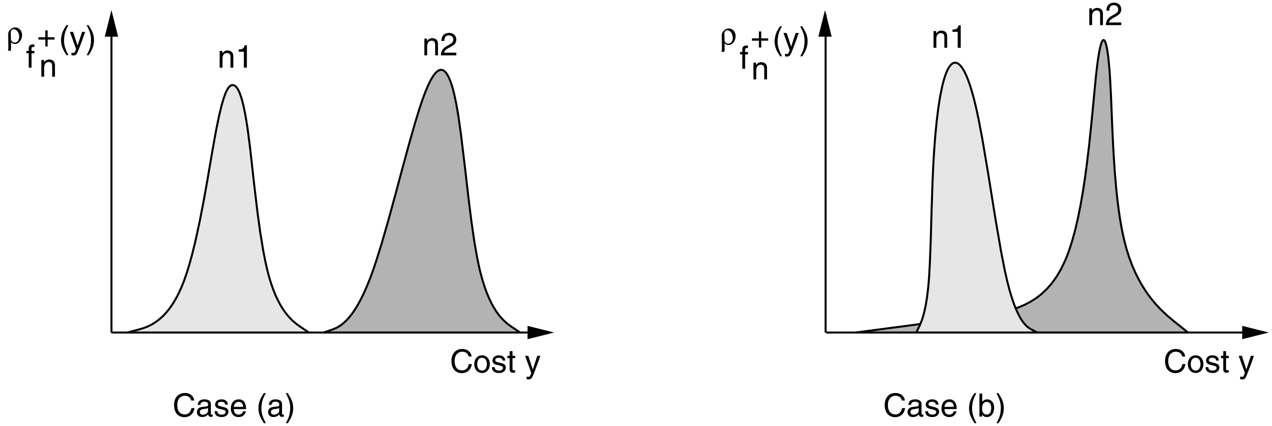


Using Information about Uncertainty of *h*

Describing the Estimation Uncertainty using Density Functions (continued)

How should an evaluation order be calculated from the density functions $\rho_{f_n^+}$ for the nodes in the OPEN list?

Possible shapes of two density functions:



Remarks:

(a) If the density functions do not overlap, the node for which the corresponding density function ρ_{f^+} has the lowest density value f_a^+ with respect to all other nodes would be selected.

(b) $f_{n_1}^+$ has the lower expected value; n_2 *has the possibility* that the cost $f_{n_2}^+$ may be lower than n_1.
An admissible algorithm would expand n_2.
It would make more sense to expand n_1 because the "$f^+(n_2) < f^+(n_1)$" event is unlikely.

➜ Due to uncertainty, costs can be overestimated or underestimated. I.e., not expanding a node in OPEN and terminating it too expensively as a result, represents a *risk*.

➜ Quantification of the risk of terminating with too high costs (= terminating too early).

Risk Measures

Defining the Order of Node Evaluations

Idea:

Estimate the risk of terminating too early using a *risk measure* R for each node in the OPEN list.

- For a given cost value C (of a goal node), the risk measure evaluates for each node n in the OPEN list to what extent C can be improved by expanding n.
- $R = R(C)$. The risk measure is a nondecreasing function of the C cost. The greater the $R(C)$ value of a node n, the greater the risk of missing an improvement of C if terminating with C without expanding n.
- $R(C)$ should use knowledge about the cost distribution for the node n, so it should be based on $\rho_{f_n^+}$.

Risk Measures

Principle of the Algorithm R*$_\delta$

Search continues until the risk value $R(C)$ of each node in the OPEN list is below a user accepted risk threshold δ.

- ➜ If a high risk is acceptable, nodes with a high risk value of $R(C)$ (= high cost reduction potential) remain unexpanded. As a result, cost underestimation becomes less likely.

  If only a small risk is acceptable, even nodes with a low risk value of $R(C)$ (= low cost reduction potential) are expanded. As a result, cost underestimation becomes more likely.

- ➜ I.e., depending on a risk threshold δ the probability of a cost underestimation (= probability of admissibility or optimality) can be controlled.

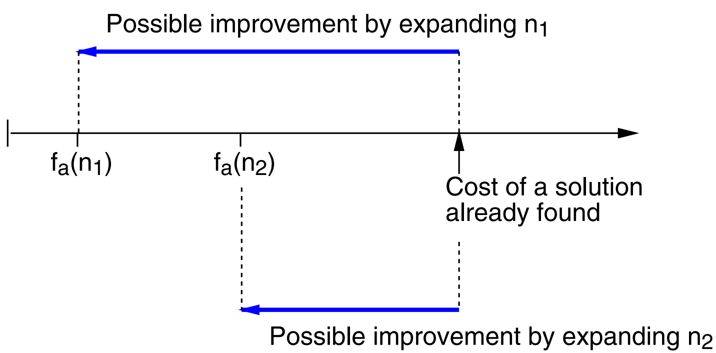
Remarks:

- The observance of this principle by the algorithm R*$_\delta$ is ensured — as shown later — by the use of certain risk functions $R(C)$.
- When a goal node with cost C is selected from OPEN, its risk function must guarantee the property $C \leq C_\delta$. Otherwise, one of the remaining nodes in OPEN can have a risk for C that is higher than δ.

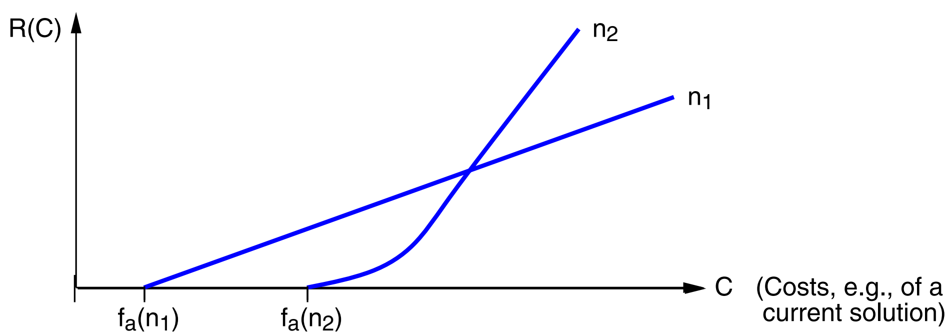
Risk Measures

Potential for Improvement to a Current Solution

Let n_1, n_2 be nodes of the OPEN list.



Example of risk functions $R(C)$ for the nodes n_1, n_2:



The nodes have different random variables $f^+_{n_1}$ and $f^+_{n_2}$ for the cost.

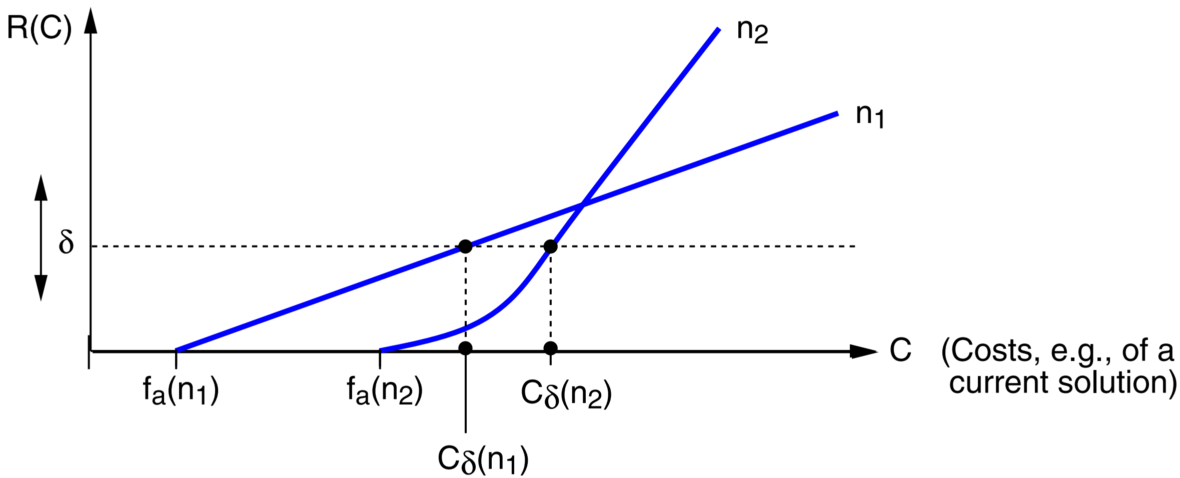
Remarks:

- The potential for improvement is a statistical quantity defined for a node n using f_n^+.
- The *evaluation* of the potential for improvement regarding given costs C is done with the help of a risk-measure $R(C)$.

Risk Measures

Risk Threshold and Cost Treshold

The risk threshold $\delta \geq 0$ defines for each node n in OPEN its cost threshold $C_\delta(n)$:



Let n_1, n_2 be nodes in OPEN. If the search was terminated with node n_2 and cost $C' = C_\delta(n_2)$, the risk $R(C')$ for n_1 would be above the risk threshold δ.

➜ R*$_\delta$ chooses the node n with the lowest cost threshold $C_\delta(n)$ in the OPEN list. In the above example, node n_1 would be preferred to node n_2.

Risk Measures

Definition 75 (Risk Measure)

Let M be the ordered set of cost values. A risk measure $R(C)$ for a node is a nondecreasing function $R : M \to [0, \infty]$ measuring the risk associated with leaving that node unexplored when terminating with a solution with cost C.

Definition 76 (Cost Threshold)

Let δ be a nonnegative real number and let $R(C)$ be the risk measure for a node n. The solution $C_\delta(n)$ to the equation $R(C) = \delta$ is called the cost threshold.

Assuming the cost of a solution path found is C, then for each node n in OPEN with $C > C_\delta(n)$ the risk of missing a better solution path is higher than risk threshold δ. These nodes should be expanded before termination.

Remarks:

- Risk measures and risk thresholds must be seen in context: not every risk threshold makes sense for a risk measure.
- Depending on the f_n^+ cost random variable of a n node, the δ risk threshold can lead to different sequences in the OPEN list.
- The cost-threshold $C_\delta(n)$ indicates how high the cost of a solution may be without exceeding the δ-risk-threshold for the node n.

Risk Measures

Definition 77 (δ-Risk-Admissibility)

An algorithm is said to be δ-risk-admissible if it always terminates with a solution cost C such that $R(C) \leq \delta$ for each node left on OPEN.

The above version of the δ-risk-admissible condition is equivalent to stating that at termination, the cost of the solution found is not greater than $C_\delta(n)$ for each n on OPEN.

Definition 78 (Algorithm R*$_\delta$)

R*$_\delta$ is a search algorithm which is identical to A* except that it chooses for expansion that node n from OPEN with the lowest cost-threshold $C_\delta(n)$.

Note that with $\delta = 0$, R*$_\delta$ is identical to A* since it is guided by the (admissible) lowest tail of the density of f, namely by $g + h_a$.

For $\delta > 0$, R*$_\delta$ may prefer a node with high f_a and narrow distribution over a node with low f_a but highly diffussed density.

Remarks:

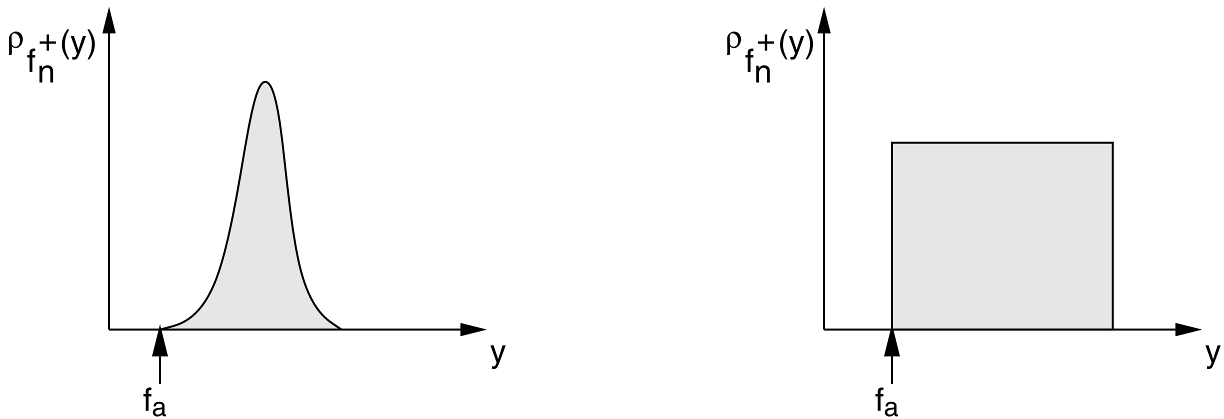
- The first definition of δ-risk-admissibility is from the perspective of risk, the second is from the perspective of cost.
- With $\delta = 0$, R*$_\delta$ is identical to A*. Justification:
  1. Computiong the cost-threshold $C_\delta(n)$ for nodes in OPEN is solving the equation $R(C) = 0$ for $\delta = 0$.
  2. $R(C) = 0$ holds for the lowest point f_a on the tail of the density of f^+.
  3. Hence, $f_a = g + h_a \leq g(n) + h^*(n)$
  4. R*$_\delta$ is guided by an admissible heuristic function and, therefore, R*$_\delta$ is admissible.
- As the δ increases, R*$_\delta$ tends to abandon admissibility.

Risk Measures

Risk Measures of Type $R(C) = \varrho[C - f^+]$

Starting point are density functions for the random variables f_n^+ of nodes n in the OPEN list.

Examples:



f_a (resp. h_a) is the smallest positive preimage of the density function $\rho_{f_n^+}$ (resp. $\rho_{h_n^*}$).

Risk Measures

Risk Measures of Type $R(C) = \varrho[C - f^+]$ (continued)

1. Worst Case Risk R_1:

$$R_1(C) = \sup_{\{y|\rho_{f_n^+}(y)>0\}} (C - y) = C - f_a = C - g - h_a$$

2. Probability of Suboptimal Termination R_2:

$$R_2(C) = P(C > f_n^+) = P(C - f_n^+ > 0) = \int_{y=-\infty}^{C} \rho_{f_n^+}(y)dy$$

3. Expected Risk R_3:

$$R_3(C) = E(\max\{C - f_n^+; 0\}) = \int_{y=-\infty}^{C} (C - y)\rho_{f_n^+}(y)dy$$

Remarks:

- The risk measures R_1 and R_3 describe costs, the risk measure R_2 describes a probability.
  - R_1:
    For the costs represented by the f_n^+ random variable, the smallest possible value is assumed. R_1 quantifies the maximum possible loss if a solution is satisfied with C costs.

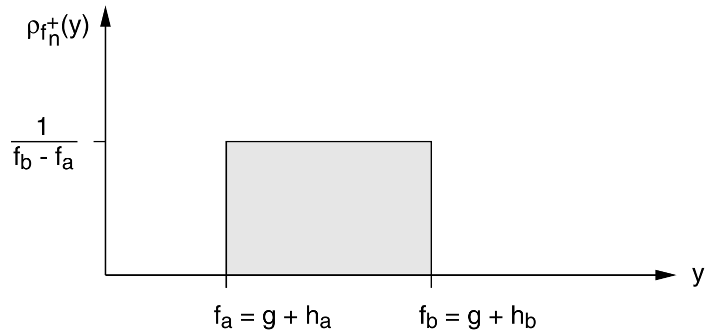
    The lowest costs are the worst case because they represent the extreme case of a missed cost reduction. The probability that the remaining costs are lower than h_a is 0.
  - R_2:
    The probability for the occurrence of the event "$C > f_n^+$" (i.e., event is a loss) is calculated if you are satisfied with a solution with C cost.
  - R_3:
    For the costs represented by the random variable f_n^+, the expected loss $E(\max\{C - f_n^+; 0\})$ is calculated if one is satisfied with a solution with costs C.

    R_3 weights the probability of the loss (R_2) with the amount of the occurring loss.

Risk Measures

Example

Let f_n^+ be uniformly distributed between an optimistic estimate f_a and a pessimistic estimate f_b (The f_a and f_b estimates depend on n, where n is in OPEN.):



Density function:

$$\rho_{f_n^+}(y) = \begin{cases} \frac{1}{f_b - f_a} & f_a \leq y \leq f_b \\ 0 & \text{else} \end{cases}$$

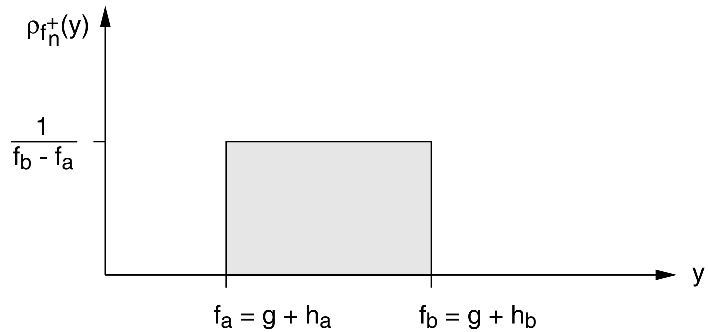
Risk measure:

$$R_1(C) = \sup_{\{y \mid \rho_{f_n^+}(y) > 0\}} (C - y) = C - f_a$$

Risk Measures

Example (continued)

Let f_n^+ be uniformly distributed between an optimistic estimate f_a and a pessimistic estimate f_b (The f_a and f_b estimates depend on n, where n is in OPEN.):



Density function:

$$\rho_{f_n^+}(y) = \begin{cases} \frac{1}{f_b - f_a} & f_a \leq y \leq f_b \\ 0 & \text{else} \end{cases}$$

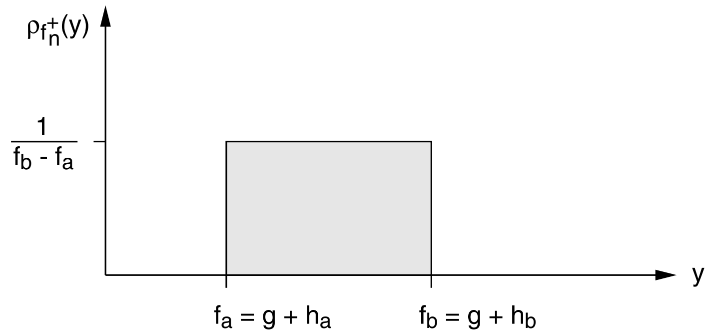
Risk measure:

$$R_2(C) = \int_{y=-\infty}^{C} \rho_{f_n^+}(y)dy = \begin{cases} 0 & C < f_a \\ \frac{(C - f_a)}{(f_b - f_a)} & f_a \leq C \leq f_b \\ 1 & f_b < C \end{cases}$$

Risk Measures

Example (continued)

Let f_n^+ be uniformly distributed between an optimistic estimate f_a and a pessimistic estimate f_b (The f_a and f_b estimates depend on n, where n is in OPEN.):



Density function:

$$\rho_{f_n^+}(y) = \begin{cases} \frac{1}{f_b - f_a} & f_a \leq y \leq f_b \\ 0 & \text{else} \end{cases}$$

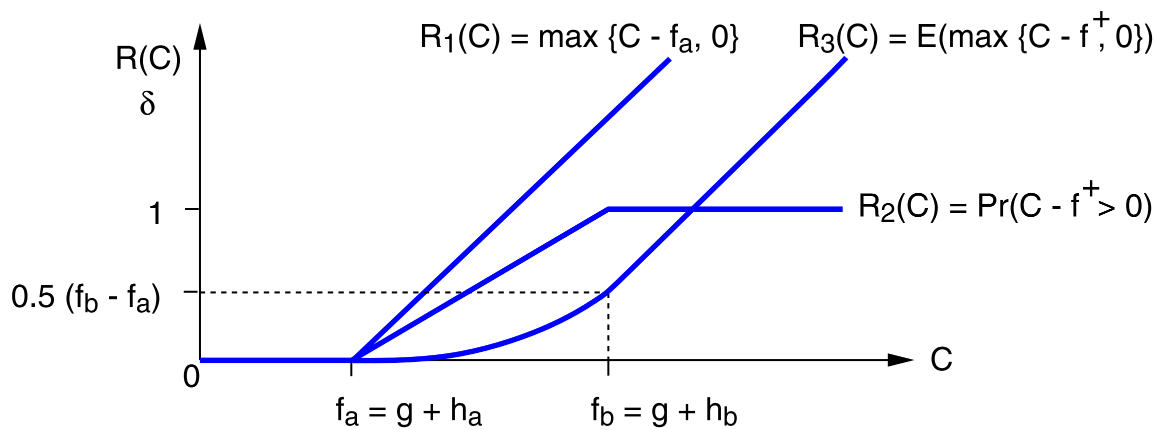
Risk measure:

$$R_3(C) = \int_{y=-\infty}^{C} (C - y)\rho_{f_n^+}(y)dy = \begin{cases} 0 & C < f_a \\ \frac{(C - f_a)^2}{2(f_b - f_a)} & f_a \leq C \leq f_b \\ C - \frac{f_a + f_b}{2} & f_b < C \end{cases}$$

Risk Measures

Example (continued)

Shape of risk measures R_1, R_2, and R_3 (f_n^+ uniformly distributed):



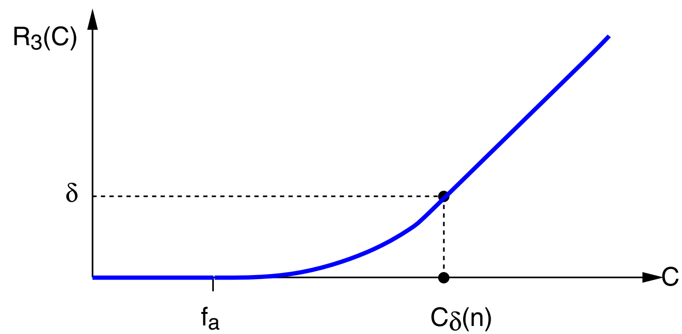
The vertical axis represents the functions $R(C)$ for the three risk measures considered.

Risk Measures

Example (continued)

Computing the cost threshold C_δ for R_3 (f_n^+ uniformly distributed):

Let δ be the user's risk tolerance of the user. For each node n in OPEN, it defines its cost threshold $C_\delta(n)$ using the equation $R(C) = \delta$.



For a node n on OPEN, $C_\delta(n)$ is computed by transforming $R_3(C_\delta) = \delta$:

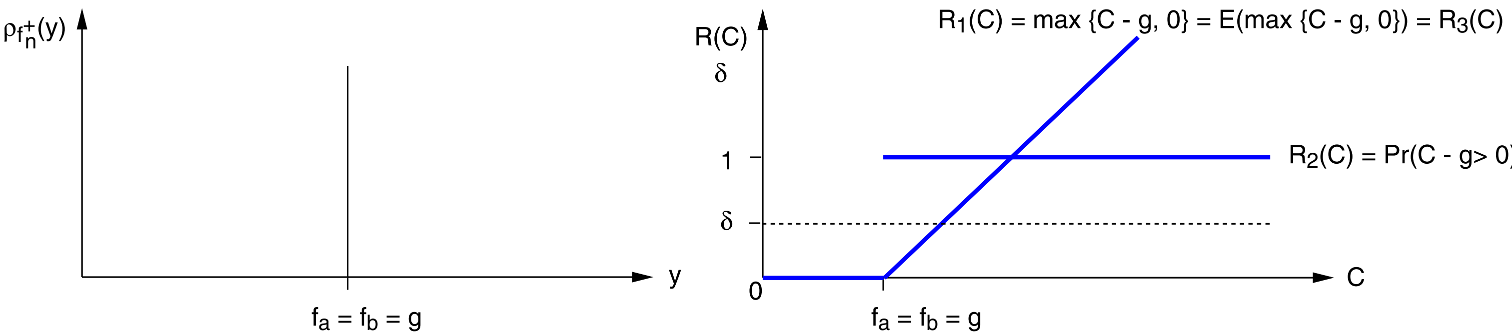
$$C_\delta(n) = \begin{cases} f_a & \delta = 0 \\ f_a + \sqrt{2 \cdot (f_b - f_a) \cdot \delta} & 0 < \delta \leq \frac{f_b - f_a}{2} \\ \delta + \frac{f_a + f_b}{2} & \frac{f_b - f_a}{2} < \delta \end{cases}$$

Risk Measures

Example (continued)

For a goal node γ in OPEN, we have $h(n) = h^*(n) = 0$. Therefore, there remains no uncertainty regarding f^+_γ.

Graphs of random variable f^+_γ for solution cost and risk measures R_1, R_2, and R_3:



Cost treshold:

$$C_\delta(\gamma) = \begin{cases} g(\gamma) + \delta & \text{for risk measure } R_1 \\ g(\gamma) & \text{for risk measure } R_2 \text{ and } \delta < 1 \\ g(\gamma) + \delta & \text{for risk measure } R_3 \end{cases}$$

Risk Measures

Theorem 79 (δ-Risk-Admissibility of R*$_\delta$)

R*$_\delta$ is δ-risk-admissible with respect to risk measures R_1, R_2, and R_3 when G is a search space graph with *Prop*(G) and $E(h^*) < \infty$ on solution paths.

Risk Measures

Theorem 79 (δ-Risk-Admissibility of R*$_\delta$)

R*$_\delta$ is δ-risk-admissible with respect to risk measures R_1, R_2, and R_3 when G is a search space graph with *Prop*(G) and $E(h^*) < \infty$ on solution paths.

Proof (sketch)

1. δ-Risk-Admissibility:

   (a) According to the previous example, it holds for the cost C of a solution path found by R*$_\delta$:

   $$C = g(\gamma) \leq C_\delta(\gamma) \qquad \text{for the risk measures } R_1, R_2, R_3.$$

   (b) Since R*$_\delta$ chooses for expansion that node n from OPEN with the lowest cost-threshold $C_\delta(n)$, δ-risk-admissibility of R*$_\delta$ follows for risk measures R_1, R_2, and R_3.

2. Completeness:

   (a) At all times OPEN contains a node n on a solution path for which $C_\delta(n)$ is finite. Obviously, $R_1(C) = \delta$ and $R_2(C) = \delta$ have a finite solution. If density $\rho_{h^*}(x)$ possesses a finite expectation $E(h^*) < \infty$ for any node on a solution path, for R_3 we have

   $$R_3(C) \geq C \cdot \big(1 - P(f^+ > C)\big) - E(f^+) \geq C - 2E(f^+) = C - 2g - 2E(h^*)$$

   (b) $C_\delta(n) \geq g(n)$ holds for each node n in OPEN since there is no risk in abandoning n after finding a solution path with cost $\leq g(n)$. A positive lower bound of the edge cost values guarantees that R*$_\delta$ can neglect nodes on solution paths only for a limited number of node expansions.

Remarks:

- Expectations can have the value ∞, e.g., for a random variable that returns values 2^n with probability 2^{-n}.

- In step 2(a) we use the fact $R_3(0) = 0$ for graphs G with nonnegative edge cost values. As the lower bound for $R_3(C)$ increases with C, there is a finite value C with $R_3(C) > \delta$. Hence, $C_\delta(n) < \infty$.

- The exact form of $\rho_{h_n^*}$ is generally unknown. For this the edge costs must have been generated by a given probabilistic model.

- Generating a good estimate for $C_\delta(n)$ is often possible. For this, the knowledge of upper and lower bounds of h_n^* together with the often reasonable assumption of a standardized distribution between them, such as an uniform distribution, an exponential distribution or a normal distribution, is sufficient.

- The principle of the ε-admissible acceleration in A*$_\epsilon$ for A* can also be applied to R*$_\delta$ and leads to the algorithm R*$_{\delta,\epsilon}$. The special version R*$_{\delta,\delta}$ is δ-risk-admissible with respect to risk measures R_1, R_2, and R_3 under the preconditions of the previous theorem.