Early Warnings of Plan Failure, Falso Positives and Envelopes: Experiments and a Model *

Paul R. Cohen, Robert St. Amant, David M. Hart

CMPSCI Technical Report 92-20

Experimental Knowledge Systems Laboratory
Department of Computer Science
University of Massachusetts
Amherst, MA 01003

Abstract

We analyze a tradeoff between early warnings of plan failures and false positives. In general, a decision rule that provides earlier warnings will also produce more false positives. Slack time envelopes are decision rules that warn of plan failures in our Phoenix system. Until now, they have been constructed according to ad hoc criteria. In this paper we show that good performance under different criteria can be achieved by slack time envelopes throughout the course of a plan, even though envelopes are very simple decision rules. We also develop a probabilistic model of plan progress, from which we derive an algorithm for constructing slack time envelopes that achieve desired tradeoffs between early warnings and falso positives.

---

* This research was supported by DARPA contract F49620-89-C-00113, by AFOSR under an Intelligent Real-time Problem Solving Initiative, AFOSR-91-0067, by ONR under a University Research Initiative grant, ONR N00014-86-K-0764, and by Texas Instruments Corporation.

1 Introduction

Underlying the judgment that a plan will not succeed is a fundamental tradeoff between the cost of an incorrect decision and the cost of evidence that might improve the decision. For concreteness, let's say a plan succeeds if a vehicle arrives at its destination by a deadline, and fails otherwise. At any point in a plan we can correctly or incorrectly predict that the plan will succeed or fail. If we predict early in the plan that it will fail, and it eventually fails, then we have a *hit*, but if the plan eventually succeeds we have a *false positive*. False positives might be expensive if they lead to replanning. In general, the false positive rate decreases over time (e.g., very few predictions made immediately before the deadline will be false positives) but the reduction in false positives must be balanced against the cost of waiting to detect failures. Ideally, we want to accurately predict failures as early as possible; in practice, we can have accuracy or early warnings but not both.

The false positive rate for a decision rule that at time *t* predicts failure will generally decrease as *t* increases. We analyze this tradeoff in several ways. First, we describe a very simple decision rule, called a *slack time envelope*, that we have used for years in the Phoenix planner (Sections 2 and 3). Then, using empirical data from Phoenix, we evaluate the false positive rate for envelopes and show that envelopes can maintain good performance throughout a plan (Section 4). An infinite number of slack time envelopes can be constructed for any plan, and the analysis in Section 4 depends on "good" envelopes constructed by hand. To be generally useful, envelopes should be constructed automatically. This requires a formal model of the tradeoff between when a failure is predicted (earlier is better) and the false positive rate of the prediction (Section 5). Finally we show how the conditional probability of a plan failure given the state of the plan can be used to construct "warning" envelopes.

2 Slack Time Envelopes

Imagine a plan that requires a vehicle to drive 10 km in 10 minutes. Figure 1 shows progress for three possible paths that the vehicle might follow, labeled A, B and C. Case A is successful: the vehicle makes rapid progress until time 3, then slows down from time 3 to time 4, then makes rapid progress until time 8, when it completes the plan. Case B is unsuccessful: progress is slow until time 4, and slower after that; and the required distance is not covered by the deadline.

The solid, heavy line is a slack time envelope for this problem. Our Phoenix planner (Cohen et al., 1989; Hart, Anderson, & Cohen, 1990) constructs such an envelope for every plan and checks at each time interval to see whether the progress of a plan is within the envelope. Case A remains within the envelope until completion; case B violates the envelope at time 6.

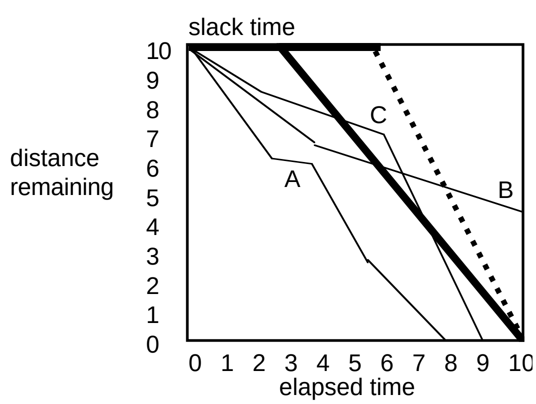

Figure 1. Illustration of envelopes.

When an envelope violation occurs, the Phoenix planner modifies or completely replaces its plan. It should not wait until the deadline has expired to begin, but should start replanning as soon as it is reasonably sure that the plan will fail. Clearly, envelopes can provide early warning of plan failure; for example, in case B, the envelope warned at time 6 that the plan would fail. The problem is that progress might pick up after an envelope violation, as shown in case C. At time 5 the envelope is violated, but by time 8, the plan is back within the envelope. If in this case the Phoenix planner abandoned its plan at time 5, it would have incurred needless replanning costs. Case C is a false positive as we defined it earlier: a plan predicted to fail that actually will succeed. Note that a different envelope, shown by the heavy dotted line, will avoid this problem. Unfortunately, it doesn't detect the true failure of case B until time 8, two minutes after the previous envelope. This illustrates the tradeoff between early warnings and false positives. (This and other concepts in the paper derive from signal detection theory, e.g., (McNicol, 1972; Coombs, Dawes, & Tversky, 1970).)

Slack time envelopes get their name from the period of no progress that they permit at the beginning of a plan. The Phoenix planner adds slack time to envelopes so that plans will have an opportunity to progress before they are abandoned for lack of progress. Until recently, this was all the justification for envelopes we could offer. In the following sections, however, we show why the simple linear form of envelopes achieves high performance, and how to select a value of slack time.

3 The Data Set

One way to evaluate slack time envelopes is to generate hundreds of plans, monitor their execution at regular intervals, and, at each interval, use an envelope to predict success or failure. We generated 1139 travel plans, or *paths*, for vehicles in our Phoenix simulation. Phoenix is based on a machine-readable map of Yellowstone National Park that includes roads, obstacles, a variety of elevations and ground covers, and other terrain features. The Phoenix planner fights simulated forest fires in this environment by surrounding the fires with fireline built by bulldozers. Envelopes in Phoenix monitor fire spread rate, fireline digging, and progress in different bulldozer tasks. The focus of this paper, however, is a simpler problem: getting from one point on the map to another by a deadline. To generate our data set, we repeatedly selected pairs of points 70 km apart as the crow flies, and asked the Phoenix planner to construct a path between each pair. Then we simulated the traversal of each path, monitoring it every 1000 simulated seconds. At each monitoring step we estimated the distance remaining to the destination. Because of obstacles, terrain, and so on, the distribution of remaining distances at a given monitoring interval was considerable (including many greater than 70 km). For example, after 5000 seconds, the mean remaining distance was about 54 km with a range of 13.6 to 79.1 km. We generated, executed and monitored 1139 paths in this manner.

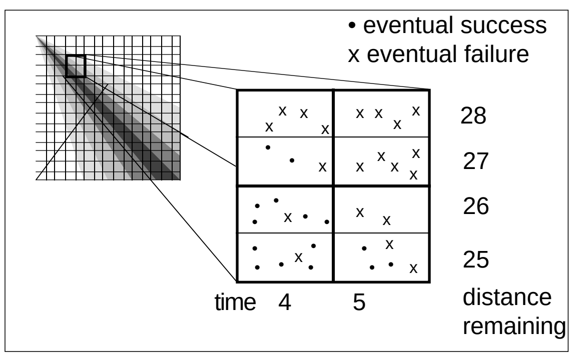

Figure 2. How we generated distributions of DR for successes and failures at each time interval.

3.1 Distributions of Eventual Successes and Failures Before the Deadline

We chose a deadline of 15,000 seconds to divide the paths into two groups: paths that reached their goals by the deadline were called *successes*, and those that did not were called *failures*. Of 1139 paths, 654 succeeded and 485 failed. We looked at each path 15 times, once every 1000 seconds, and recorded an estimate of the number of "distance units" remaining to the goal. For a variety of reasons, a distance unit is 2 km, so the distance remaining to the goal, abbreviated DR, is 35 at the beginning of the plan and zero for successful paths at the end of the plan. Henceforth, we use "time x" as shorthand for "x thousand seconds elapsed." For example, in Figure 2, at time 4, all the paths with DR = 28 are failures; at time 5, all the paths with $26 \le DR \le 28$ are failures; but at time 5, DR = 25, three paths are successes and two are failures.

We plotted frequency polygons for DR for successes and failures at each of the 15 time intervals. Figure 3 shows the distribution of successes and failures at time 5. Note that most failures still have a long way to travel at time 5: the bulk of the distribution lies to the right of DR = 30 (the mean DR for failures at time 3 is 33). The distribution of successes, however, is made up of paths with relatively short remaining distances to the goal (mean DR = 17).

3.2 Empirical Hit Rates and False Positive Rates for DR Thresholds

Let's predict that a path will fail to reach its goal by its deadline if, at time 5, the remaining distance to the goal is 30 or more, that is, the *threshold* DR ≥ 30. The dark shaded area in Figure 3 represents false positive

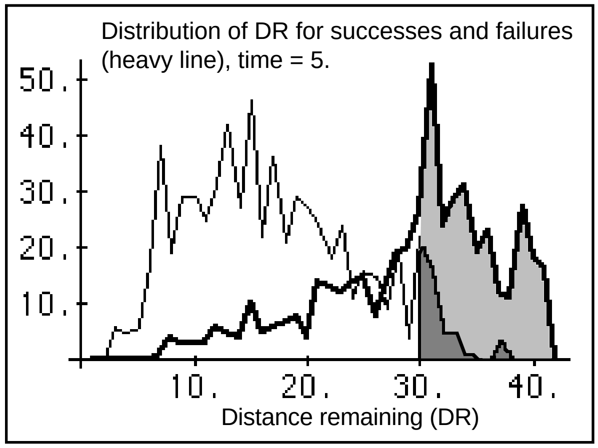

Figure 3. Frequency polygons for DR at time 5.

errors, paths we predict will fail but that eventually succeed. Of the 654 paths that eventually succeed, 37 lie in the dark shaded area; the probability of a false positive is therefore 37/654 = 0.056. The light shaded area represents hits, paths that are predicted to fail and that actually do fail. The probability of a hit is 261/485 = 0.538. The ratio of the probabilities is 9.60.

At time 5, the threshold DR ≥ 30 seems pretty good because the ratio of hit probability to false positive probability is high, but we cannot say it is the *best* threshold unless we know the relative values of hits and false positives.

This is just part of the analysis of our data set. In particular, we haven't shown our analyses of success and failure distributions at other times; nor hit and false positive probabilities for different DR thresholds at different times; nor success and failure distributions for stricter or more lenient deadlines. We can summarize these analyses as follows: At later time intervals, the success distribution is increasingly right skewed, with most of its mass around low values of DR. At intermediate time intervals (e.g., time = 7) the failure distribution is roughly uniform. Later, it is right skewed like the success distribution, but with more mass in its tail than the success distribution. Shifting the DR threshold to the right decreases both hits and false positives, though false positives decrease faster (as in Fig. 5, only more so at later time intervals). These patterns hold for stricter and more lenient deadlines; the main effect of a stricter deadline is to reduce the number of successes. The following evaluations of envelopes are based on the 15,000 second deadline illustrated above because it produces a nearly even split between successes (654, total) and failures (485, total).

4. Evaluation of Slack Time Envelopes

Slack time envelopes are decision rules for predicting whether paths will succeed or fail. As illustrated in Figure 1, if the path is within the boundary of an envelope at a particular time, then we predict success, otherwise we predict failure. Each point on an envelope boundary specifies a DR threshold for a particular time, and so has an associated hit rate and false positive rate. In this section we use slack time envelopes to predict whether paths in our data set, discussed above, will succeed or fail. We evaluate the predictive performance of slack time envelopes according to this criterion: An envelope should provide performance approaching optimal throughout a plan.

This depends of course on our definition of optimal. Consider a decision rule based on distance remaining (dr) to the goal at time t:

$$\textit{If } l(dr, t) > \beta(t) \textit{, then predict plan failure}$$

$$\text{where } l(dr, t) = \frac{\mathbf{Pr}(dr \mid \textit{plan fails}, t)}{\mathbf{Pr}(dr \mid \textit{plan succeeds}, t)}$$

Intuitively, we have an observation of dr at time t, and we must decide whether this observation has been produced by an eventual success or failure. We base our decision on whether the likelihood is greater than the threshold $\beta(t)$. A basic result from signal detection theory is that the utility of this decision is maximized if

$$\beta(t) = \frac{\mathbf{Pr}(\textit{plan succeeds}, t)}{\mathbf{Pr}(\textit{plan fails}, t)} (\textit{Payoff}(t))$$

$$\text{where } \textit{Payoff}(t) = \frac{\textit{Val}(\textit{correct rej}, t) + \textit{Cost}(\textit{false pos}, t)}{\textit{Val}(\textit{hit}, t) + \textit{Cost}(\textit{miss}, t)}$$

In the simplest case, $\beta(t)$ is constant over the course of a plan. A more realistic assessment requires analysis of the terms in $\beta(t)$. The first term, the prior probabilities, decreases with time, as plans begin to succeed. The second term *Payoff(t)* determines the relative importance of hits, false positives, correct rejections, and misses. The value of correctly predicting a plan failure decreases over time; early warnings are worth more. At the same time the cost of a false positive increases over time; if we are going to unnecessarily abandon a plan, it is better to do so early in the plan than later when we have invested a lot of time in the plan. It is more difficult to assess the value of a correct rejection and the cost of a miss, but if we assume they are constant relative to the other parameters, then the value of the second term in *Payoff(t)* increases over time. We consider the cases in which *Payoff(t)* is constant and also in which *Payoff(t)* increases linearly with time.

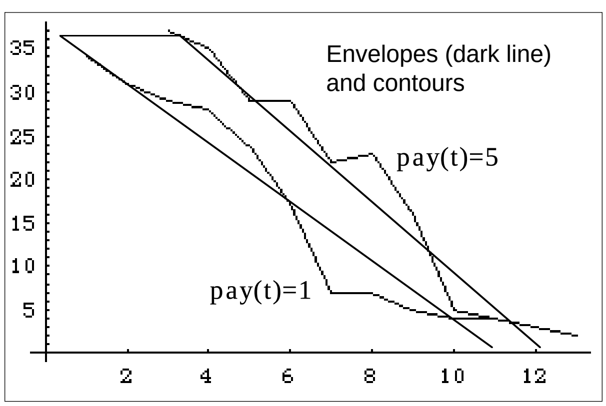

Figure 4. Hand-constructed slack time envelopes superimposed on constant *Payoff(t)* contours.

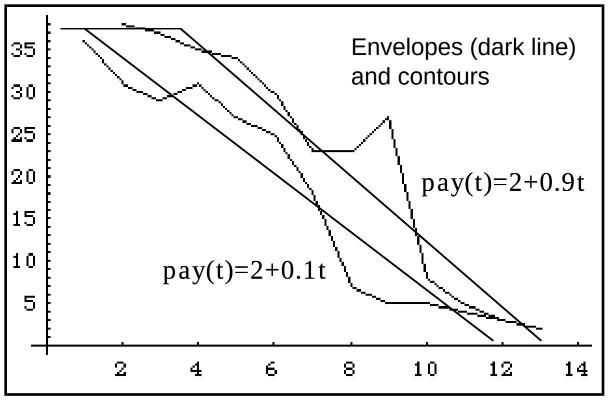

Figure 5. Slack time envelopes on linear *Payoff(t)* contours.

4.1. Comparing Slack Time Envelopes with Empirical Utility Contours

Because envelopes are just straight lines, it is unclear whether they can satisfy the optimal performance criterion. In particular, for constant or linear payoff functions, the DR threshold required to maintain a constant ratio of hit probability to false positive probability might not change linearly over time. To find out, we calculated utility *contours* from the empirical data for different *Payoff(t)*, as shown in Figures 4 and 5. A contour represents a fixed *Payoff(t)* function; each point on a contour is the DR threshold (y axis) that is required at a particular time (x axis) to ensure that the utility of the decision is maximized. In Figure 4 we let *Payoff(t)* be constant at 1 and 5, and in Figure 5 we let *Payoff(t)* vary as a function of time.

An important characteristic of these contours is that they require high DR thresholds for the first few time intervals, but then gradually smaller thresholds for later time intervals. Utility contours are roughly linear, which suggests that a slack-time envelope, fit to one of these contours, ought to provide performance approaching optimal, given our payoff function. For our data set, at least, Figures 4 and 5 tells us that an envelope can be constructed to satisfy our performance criterion. We will formalize this result in the next section.

5. Constructing Slack Time Envelopes: How Much Slack?

Our focus now turns to the task of constructing slack time envelopes. We assume that the end points of the envelope are the distance to the goal and the deadline, so the only parameter is how much slack time to allow. Next we present a model that predicts utility for different values of slack time. The model also predicts the *early warning premium* for values of slack time. Early warning premiums accrue when, by constructing a tight envelope with little slack time, we detect failures earlier than we would with a looser envelope. Empirically, early warning premiums come at the expense of false positives. We assess a cost for each hit proportional to the time interval in which it is detected; this places a premium on early hits. We assess a constant cost for each false positive. This is described further in the following sections.

5.1. A Probabilistic Model of Progress

If we know the distributions of distance remaining (DR) for successes at each time interval (e.g., those in Figure 3) then we can predict the false positive rate for a given DR threshold. A simple model of the distribution of DR begins with the assumption that in each time interval a vehicle can progress at its maximum rate c with probability p, or makes no progress at all with probability q = 1 - p. Then the distribution of progress is binomial, as shown in Table 1: the probability of having made r units progress by time n is just the binomial probability $\binom{n}{r} p^r q^{(n-r)}$.

For example, the probability of one unit progress by time 4 is $4pq^3$ because there are four ways to achieve this result, each with probability pq^3: we could make no progress until time 3 (with probability q^3) and then progress at the maximum rate for one time unit (total probability, pq^3). Or we could make one

Progress \ Time	1	2	3	4	5
0c	q	q^2	q^3	q^4	q^5
1c	p	pq	$3pq^2$	$4pq^3$	$5pq^4$
2c		p^2	$3p^2q$	$6p^2q^2$	$10p^2q^3$
3c			p^3	$4p^3q$	$10p^3q^2$
4c				p^4	$5p^4q$
5c					p^5

Table 1. Progress in each time interval follows a binomial distribution.

unit of progress by time 3 (with probability $3pq^2$) and then make no progress for the remaining time unit (total probability $3pq^3$). The sum of these options is $4pq^3$.

The expected progress after N time units is cNp and the variance is cNpq. If $p = q = .5$ then the distributions of progress in each time interval are symmetric. Otherwise the mass of the distribution at time N tends toward cN (if $p > q$) or zero (if $q > p$). Important characteristics of this model are that progress is linear and variance changes linearly with time.

5.2 Utility Contours Using the Model

This model explains the shape of utility contours and slack time envelopes, and it predicts the probability of false positives for a given envelope. Let us elaborate the model a little: Our goal is to travel some distance D_g by a deadline time T_g. At any time t, we can assess the progress that has been made, D(t), and the progress that remains to be made, DR(t); and the time remaining, $TR = T_g - t$. A success is defined as $D(t) \geq Dg$ and $t \leq Tg$. The conditional probability of a success given DR(t), Dg and Tg is:

$$\mathbf{Pr}(success \mid DR(t)) = \sum_{r = DR(t)}^{TR} \binom{TR}{r} p^r q^{(TR - r)}$$

A similar equation holds for the conditional probability of a failure. If *Payoff(t)* is for example constant, this means means that the ratio of these conditional probabilities must be constant as well. Now imagine that we have DR(t) distance remaining at time t and we extrapolate forward TR time units to the deadline. At this point we have a binomial distribution with N = TR, divided into a portion below the DR=0 line (the successes, those cases that have arrived by the deadline) and a portion above the line (the failures.) The ratio of the areas of the two portions gives us the ratio of the conditional probabilities. If we want to find at each time the distance for which this ratio is constant, we plot a constant z-score for distributions with N ranging from T_g to 0.

Figure 6 shows contours for constant *Payoff(t)*. Contours for comparable linear *Payoff(t)* are very similar, with identical slack times, but more pronounced curve. To generate the figure we assumed $D_g = 25$, $T_g = 50$, p = .5, and c = 1, and applied the above analysis to get conditional probabilities of success and failure for every value of t.

Imagine that a vehicle has made 10 units of progress at time 25, that is, DR(25) = 15, illustrated by the large dot near the center of Figure 6. Because this dot lies on the contour labelled *Payoff(t)* = 5, we know that Pr(failure | DR(25) = 15) / Pr(success | DR(25) =15) = 5. If the vehicle makes no progress for another five time units, then the dot would lie to the right of the contour labeled *Payoff(t)* = 43, so the probability ratio is much higher.

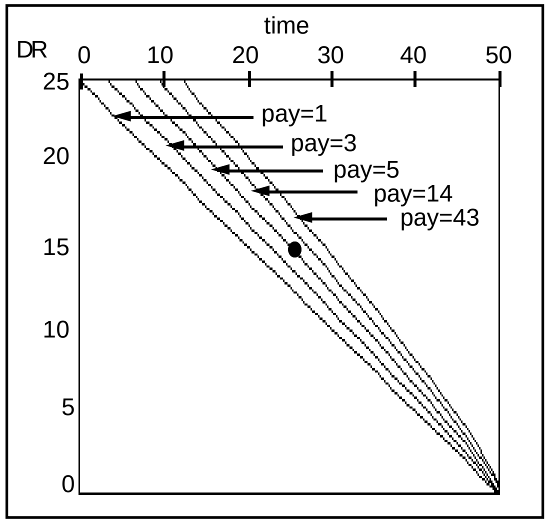

Figure 6. Contours of constant payoff from each point in the space.

These contours vary as $\sqrt{t}$. At the scale on which we monitor, linear envelopes provide a good approximation of the contours, as long as the envelope boundaries have the right slope, that is, if they are constructed with the right amount of slack time. Note, too, that Figure 6 justifies the use of slack time in envelopes: The contours associated with high payoffs (and thus high ratio of hit probability to false positive probability) allow a period of no progress at the beginning of the plan.

5.3 Setting Slack Time

A slack time envelope is just a pair of lines, one representing the period in which no progress is required—the slack time—and another connecting the end of the first to the deadline, as shown in Figures 1 and 6. Slack time is the only parameter in slack time envelopes, but we must still show how to set it.

We desire a balance of false positives against early warning premiums. We have not yet derived from our binomial model a closed-form expression for the expected number of false positives and early warnings, but we have an algorithm that produces these expectations for a given value of slack time, if we assume that $D_g = .5\ T_g$:

For each possible value of DR, dr_i:

a. calculate t_e, the time at which the envelope boundary will be crossed, given dr_i; for example, in Figure 1, when $dr_i = 5$ and $t>8$, the solid envelope boundary is crossed, so for $dr_i = 5$, $t_e = 8$.

b. use the binomial model to calculate p_e, the probability of reaching t_e; for example if $dr_i = 3$ and $t_e = 5$, Table 1 tells us that $p_e = 10p^2q^3$.

c. use the model to find the probability of a false positive, $p_{fp} = Pr(success \mid DR(t_e) = dr_i)$.

d. $p_e \times p_{fp}$ is the probability of a false positive for this value of dr_i

e. $p_e \times (T_g - t_e)$ is the expected early warning premium for this value of dr_i.

$T_g - t_e$ is the time that remains before the deadline at the envelope boundary at dr_i; this is why $T_g - t_e$ is called the early warning premium. The expected early warning premium for a value of dr_i is just $T_g - t_e$ times the probability of crossing the envelope boundary. The *mean expected early warning premium* is the mean over all values of dr_i of $p_e(T_g - t_e)$. We expect it to have higher values for lower slack times, because the envelope boundaries for low slack times are further from the deadline. The *mean probability of false positives* is obtained by summing $p_e\ p_{fp}$ for all values of dr_i and dividing by the number of these values. We expect it to rise, also, as slack time decreases, as suggested by the contours in Figure 6.

With a table of values for the mean probability of false positives and the mean expected early warning premium, and utilities for early warning and false positives, we can make a rational decision about slack time.

6 Conclusion

Although we rely heavily on slack time envelopes in the Phoenix planner, we have always constructed them by heuristic criteria, and we did not know how to evaluate their performance. In this paper we showed that high performance can be achieved by hand-constructed slack time envelopes, and we presented a probabilistic model of progress, from which we derived a method for automatically constructing slack time envelopes that balance the benefits of early warnings against the costs of false positives.

Other work has been done in this area, e.g., (Miller, 1989) constructs an execution monitoring profile of acceptable ranges of sensor values for a mobile robot (this profile is also called an "envelope"). If, during plan execution, a sensor value exceeds the envelope boundaries, a reflex is triggered to adjust the robot's behavior in such a way that the sensor readings return to the acceptable range. (Sanborn and Hendler, 1988) have used monitoring and projection in a simulated robot that tries to cross a busy street. The robot has a basic street-crossing plan, but monitors oncoming traffic and predicts possible collision points which trigger reactive avoidance actions. Our contribution has been to cast the problem in probabilistic terms and to develop a framework for evaluation. We are currently extending our work to other models of progress and different, more complex domains. A technical report covering this work in more detail is in preparation.

Acknowledgements

This research is supported by DARPA under contract #F49620-89-C-00113, by AFOSR under the Intelligent Real-time Problem Solving Initiative, contract #AFOSR-91-0067, and by ONR under a University Research Initiative grant, ONR #N00014-86-K-0764, and by Texas Instruments Corporation. We wish to thank Eric Hansen and Cynthia Loiselle for many thoughtful comments on drafts of this paper. The US Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

References

Cohen, P.R.; Greenberg, M.L.; Hart, D.M.; and Howe, A.E., 1989. Trial by Fire: Understanding the Design Requirements for Agents in Complex Environments. *AI Magazine*, 10(3):32-48.

Coombs, C.; Dawes, R.; and Tversky, A., 1970. *Mathematical Psychology: An Elementary Introduction.* Ch. 6. The Theory of Signal Detectability. Prentice Hall.

Hart, D.M.; Anderson, S.D.; and Cohen, P.R., 1990. Envelopes as a Vehicle for Improving the Efficiency of Plan Execution. In *Proceedings of the Workshop on Innovative Approaches to Planning, Scheduling, and Control*. K. Sycara (Ed.) San Mateo, CA.: Morgan-Kaufmann, Inc. Pp. 71 – 76.

McNicol, D., 1972. *A Primer of Signal Detection Theory*. George Allen and Unwin, Ltd.

Miller, D.P., 1989. Execution Monitoring for a Mobile Robot System. In SPIE Vol. 1196 *Intelligent Control and Adaptive Systems*. Pp. 36-43.

Sanborn, J.C.; and Hendler, J.A., 1988. Dynamic Reaction: Controlling Behavior in Dynamic Domains. *International Journal of Artificial Intelligence in Engineering*, 3(2).