Extending Classical Planning to Real-World Execution with Machine Learning

Gerald Dejong
Computer Science/Beckman Institute
405 N Mathews, Urbana IL 61801

Scott Bennett
Systems Research and Applications Corporation
2000 15th St North, Arlington VA 22201

Abstract

In previous work (Bennett 1993 DeJong and Bennetl 1993) we proposed a machine learning approach called *permissive planning* to extend classical planning into the realm of real world plan execution Our prior results have been favorable but empirical (Bennetl and DeJong 1991) Here we examine the analytic foundations of our empirical success We advance a formal account of real-world planning adequacy We prove that permissive planning does what it claims to do it probabilistically achieves adequate real-world performance or guarantees that no adequate real-world planning behavior is possible within the flexibility allowed We prove thai the approach scales tractably We prove that restrictions are necessary without them permissive planning is impossible We also show how these restrictions can be quite naturally met through schema based planning and explanation-based learning

1 Introduction

Real-world execution of classical plans can be problematic Small but unavoidable discrepancies between the system s representations and the real world often conspire to cause real world failure We adopt a model of real world planning adequacy Next we discuss *planner bias* an inescapable facet of classical planning We show that through adjustment of planner bias real world adequacy can be achieved The adjustment process called *permissive planning* is empirically guided Thus it is a form of machine learning In permissive planning the planners domain theory and projection capabilities remain unchanged Through bias adjustment the real-world manifestation of produced plans are brought into line with projection according to the micro world

1 1 Planning and Real-World Adequacy

A classical planner can be viewed as accreting constraints until the set of action sequences consistent with the constraints are all believed to solve the planning problem We term a classical planner together with its declarative world model (its operator definitions etc) a *planning s\item A planning domain* is a source ot planning problems together with some real world interface We model the planning problems as an unlimited sequence whose elements are randomly generated by the world according to some fixed but unknown distribution over a universe of well-formed problems The real-world interface must minimally be able to execute a plan and report whether the goal of the planning problem is satisfied in the resulting stale

This research was sponsored in part by the Office of Naval Research under grant N00014-94-1 -0684

The planning system is applied to domain problems in sequence For each problem the planner expends up to some fixed resources and either offers a solution or choose not to offer a solution (by outputting the special symbol Λ)

We say a plan *ISolves* (solves according to the system s own Internal model of the world) a planning problem if the initial state projected through the plan satisfies the problems goal The plan *ESolves* (solves according to the External world) a planning problem if the real-world state that results when the plan is executed from the problems initial state satisfies the problems goal Likewise a *planning system* ISolves (or ESolves) a problem il the plan it produces ISolves (or ESolves) the problem

Definition The *adequacy* of a planning system P over a universe of problems U randomly sampled according to a distribution *D* is $A_{UD}(P) =$

$$\sum_{\iota \in U} \Pr_D(\iota) \bullet \begin{cases} \alpha \text{ if P ISolves } \iota \text{ and P ESolves } \iota \\ \beta \text{ if P does not ISolve } \iota \\ \gamma \text{ if P ISolves } \iota \text{ but does not ESolve } \iota \end{cases} \quad (1)$$

where α β and γ are constants with $\gamma \le \beta \le 0 < \alpha$ and $\Pr_D(\iota)$ is the probability of occurrence of problem ι according to distribution *D*

Each contribution is weighted by the probability of encountering the problem according to the distribution Adequacy improves if an offered plan actually works in the real world producing a bad plan cannot be better (and may be worse) than producing no plan at all

1 2 Planner Bias Space

Planner bias is the preference (exhibited bv all classical planners) for one solution over others When presented with a problem the planners abilities are in general capable of constructing many quite different solutions For breakfast one might be capable of making pancakes crepes Belgian waffles hot oatmeal cold raisin bran sausage and eggs gravy and biscuits eggs Benedict etc We call this the *competence set* of the planner No planner will construct them all planning activity halts when the first solution is found We call this preferred element of the competence set *the performance item*

We wish to examine the effect of altering the bias of a planner The bias of a planner is just a selection function for any planning problem, the bias selects either an element of the competence set or Λ The competence set of a problem remains the same if the bias is changed but a different performance item may be selected Now consider a space of possible biases The notion though abstract is well formed A bias space is simply a collection of possible selection functions

Given a bias space and a planning system it is entirely possible (indeed quite likely) that for a given domain some biases will result in a higher adequacy than others By systematically searching the bias space we might hope to achieve acceptable real-world behavior We call such a search *Permissive Planning*

2 Ideal Permissive Planning is Impossible

In ideal permissive planning the bias space is unrestricted There are no *a prion* constraints upon the individual biases that make up the space and the bias space may be infinite

Definition *ideal permissive planning* is a recursive procedure which when given

- a planner a bias space adequacy values α β γ
- a threshold of acceptable adequacy T_a between a and β
- a source of planning problems

produces a new planner through bias accommodation which has the following properties

- if the space contains no bias of adequacy $\geq T_a$ the new planner refuses to offer solutions to any further problem
- otherwise the bias with the highest adequacy is adopted

The recursiveness of the procedure requires that the algorithm always halt Therefore only a finite number of example problems can be examined before either selecting the best bias or guaranteeing that no acceptable bias exists Under these conditions the following theorem holds

Theorem 1 Ideal permissive planning cannot be realized

The proof follows directly from Lemma 1 concerning *simplified ideal permissive planning* a greatly restricted version of the ideal formulation Here the bias space contains e*xactly two* biases Furthermore T_a=0 and $\gamma = \beta$ After attempting to solve a finite number of world problems the better of the two biases must be selected Problems for which the planner offers no solution are counted as failures

Lemma 1 The simplified formulation of ideal permissive planning cannot be realized

Proof Under these conditions the biases are independent Bernoulli random variables We can know their characteristics only by sampling world problems and observing success or failure This is precisely the discrete ume minimax bandit problem with two Bernoulli arms (Presman and Sonin 1990) as the horizon *N* grows without bound A permissive planning algorithm is any decision proceduTe which after a finite sampling of the two biases selects once and for all the better of the two This corresponds to a regret function that is asymptotically flat in *N* An important result from the bandit literature is that the optimal regret function grows at least as $\sqrt{N}$ Since no regret function can under these conditions be asymptotically flat no algorithm can realize simplified permissive planning *QED*

Any ideal permissive planning algorithm is necessarily also an algorithm for simplified ideal permissive planning so Theorem 1 holds

3 Restricted Permissive Planning Is Tractable

This negative result is not the last word on permissive planning We now show that a particularly simple kind of permissive planning, which we call restricted permissive planning is tractable Instead of guaranteed optimality we will content ourselves with *any* bias above the specified threshold We also allow a small non-zero adequacy tolerance parameter E that establishes an indifference region about the chosen adequacy threshold If the adequacy of the best bias in the space falls within $\pm\varepsilon$ of the threshold we are indifferent to the outcome of permissive planning it may either adopt a bias or decide there is no adequate bias and choose not to solve further problems Finally restricted permissive planning need only succeed with a high probability More precisely

Definition *restricted permissive planning* is a recursive procedure which when given

- a planner a bias space adequacy values α, β γ
- a threshold of acceptable adequacy T_a between a and β
- a source of planning problems
- two parameters £ and 6 each between 0 and 1

produces a new planner through bias accommodation which with probability of at least $1-\delta$ has the following properties

- if the bias space contains any bias of adequacy $\geq T_a+\varepsilon$ some bias with adequacy of at least T_a E will be adopted
- if the space contains no bias of adequacy $\geq T_a$ ε, the new planner refuses to offer solutions to any further problem
- if the best bias in the space falls within the indifference range of $T_a \pm \varepsilon$ then the planner may either adopt a bias of adequacy $\geq T_a-\varepsilon$, or refuse to offer solutions to further problems

The ε/δ arrangement is borrowed from the PAC literature (Valiant 1984) and earlier employed in (Bechhofer 1954) From this point in the paper on when we mention adequacy we will understand it to mean $1-\delta$ probabilistic attainment of an ε-adequacy threshold

3.1 Schema-Based Planners and their Bias Spaces

In a schema (or skeletal) planner planning techniques, are encoded as generalized solution patterns called schemata The system solves problems by retrieving an a general solution and instantiating it to fit the particular problem falling back on a search planner only if no schemata apply A schema can be viewed as a mechanism for supplying a set of constraints which depend in part on the problem s initial slate and goal

Definition a *schema* SC is a set of constant constraints $\{C_{SC}\}$ on action sequences together with a function F_{Sc} $P \rightarrow C^* \cup \Lambda$ that maps a planning problem into either a set of additional constraints or the special symbol Λ If F_{sC} yields a constraint set for a problem $\{C_{sc}\} \cup F_{sc}$ denotes only action sequences that ISolve the problem

The symbol Λ is interpreted as indicating that the schema s preconditions are not satisfied by the problem The collection of all the system s schemata is its schema library The *bias* of a schema-based planner is the particular library u possesses Any change to a schema shifts the bias even if the replacement schema is only slightly different from the original We wish to consider biases composed of schemata acquired through EBL over problems sampled from the planners distribution

As an aside problems often naturally cluster into problem classes The potential goals and common initial slale leatures for getting to a distant city are significantly different than those involved in washing clothes We will assume lhai problem classes are pre-existing and that it is easy for the planner to tell which class a problem is from This is not to say that it is necessarily easy to know which schemata apply to a problem Schema applicability is determined by the satisfiability of each schema s preconditions From a formal point of view all ot our results, hold if we view all problems as coming from a single class But problem class is an important facet of schema planning and we prefer not to ignore it If problem classes are used however we insist that adequacy be judged for each class separately it would be misleading to allow a planner to balance *inferior* performance on difficult problem classes against infallible performance on easy ones

3.2 Honest EBL

When given a problem we will require that the underlying EBL system produce at random one of the schemata from the set of possible EBL-acquirable schemata Furthermore the generation process must be fair We require that any schema that can in principle be produced will be produced from the problem with no less than some finite minimal probability p_{min} That is we exclude EBL systems which can hide schemata behind vanishingly small construction probabilities From this it follows that a problem can lead to only a finite number of schemata in fact no more than $1/p_{min}$ This is usually the case in EBL applications Note there is no upper bound on the number of different schemata that might be constructed from the domains problems and that the bias space may include an infinite number of distinct biases

We also insist that a schema s preconditions not be artificially specific the preconditions must accurately portray the capabilities of the schema s body For example a schema sufficient to ISolve general grocery shopping problems is not allowed to claim only to ISolve problems of acquiring dairy products More formally there cannot exist within the precondition language a more-general goal pattern nor a less constraining initial state requirement that accurately captures the applicability of the schema We will call such a schema honest since it can commit no sin of omission in advertising its applicability We will say an EBL system is honest if u produces only honest schemata

Our EBL system then is assumed to have the following characteristics We provide it with some resource bound R and allow access to a domains problems When given a problem EBL expends at most R resources to produce *a* schema This succeeds with some probability call it *p* With probability 1 *p* the EBL system produces Λ indicating that no schema will be forthcoming When a schema is produced each of the set of acquirable schemdta appears with probability no less than p_{min} Only schemata that honestly characterize their preconditions are produced The EBL behavior over a domain can be summarized by p_{min} and *p* Note that increasing R serves to increase p

3.3 Conformable Schema Application

Now we turn to the planner in which the EBL system is embedded This is the consumer of the EBL-acquired schemata We will say that a schema based planner is conformable if a schema can be applied to a problem only if the problem s goals require all of the schema's goal pattern To illustrate consider a schema for achieving the conjunctive goal pattern $ON(?x\ ?y) \wedge ON(?y\ ?z)$ This schema is sufficient to achieve the simple problem goal ON(Blk1 Blk3) but such a use would not be conformable since a portion of the schema s goal pattern is not matched by the problem goal It might seem more desirable to apply an overly specific schema than to fail *to produce* a solution altogether However a schema selection failure conveys important information to the learning component about the abilities of the schema based planner This information is obscured if schemata can be employed in a non-conformable (and incidentally often inefficient) manner

We can now stale the following lemmas

Lemma 2 The probability that a schema which has adequacy T_a ESolves a random problem of the appropriate problem class is at least $\frac{T_a - \beta}{\alpha - \beta}$

Proof Suppose we give the schema k problems of which r are both ISolved and ESolved p are nor attempted and q are ISolved but not ESolved By hypothesis the schema meets or exceeds the adequacy threshold T_a From (1)

$$\frac{r\alpha + p\beta + q\gamma}{k} \geq T_a \qquad (2) \quad \text{or equivalently}$$

$$r \geq \frac{kT_a - q\gamma - p\beta}{\alpha} \qquad (3)$$

Let us examine a lower bound on the ratio $\frac{r}{k}$ Consider how k problems can be apportioned among r q, and p From (3) we see that for a given T_a r is minimized when q=0 (since $\gamma \leq \beta \leq 0 < \alpha$) If q=0 then p=k-r Substituting these into (2) and solving for r we obtain $\frac{r}{k} \geq \frac{T_a - \beta}{\alpha - \beta}$ This ratio is just the proportion of problems which are both ESolved and ISolved to the total number of problems QED

Lemma 3 The probability that a schema which has adequacy T_a ISolves a random problem of the appropriate problem class is at least $\frac{T_a - \beta}{\alpha - \beta}$

(This is established in the proof of Lemma 2)

Lemma 4 Given an honest EBL schema acquisition system L and a conformable schema application planner A it is the case that if a planning problem P can be ISolved by A using a schema SC then SC is acquirable by L from P

Proof Sketch We can think about each schema as embodying a planning strategy that can be applied directly without recourse to searching the constraint tree Acquiring planning schemata through EBL can be viewed as the process of 1) observing an instance that illustrates an unknown strategy 2) explaining why the instance works 3) abstracting away unneeded and easily redenvable details of the example to yield a general characterization of the illustrated strategy If a planning strategy suffices to solve a problem, men some from-scratch plan can be constructed (by the planner underlying the EBL component) which is an example of the strategy This plan suitably explained can therefore be generalized back into the original strategy Of course, alternate from-scratch solutions might lead to other strategies and therefore other schemata But among the schemata acquirable from the planning problem must be the originally postulated schema The only difficulty in transforming this informal argument into a proof is to formally insure a tight fit between the problem and the strategy the problem solution must necessarily illustrate all facets of the strategy This property is enforced by the honesty and conformability requirements

Lemma 5 If a planner contains a schema, SC acquirable through honest EBL which has adequacy T_a or higher for a problem class then the probability that EBL will in fact construct SC from an arbitrary problem of the appropriate class is at least $\frac{p_{min}(T_a - \beta)}{\alpha - \beta}$ where α and β are adequacy parameters from Definition 1 and p_{min} is minimum schema probability factor of the EBL system

Proof By hypothesis the planner contains a schema of adequacy at least T_a Call that schema SC From Lemma 3 the expected ISolve solution rate of SC over relevant problems is at least $\frac{T_a - \beta}{\alpha - \beta}$ From Lemma 4 we know that each of the ISolvable problems could in principle give rise to the acquisition of SC through honest EBL Thus the expected fraction of problems of the appropriate class that could in principle give rise to SC is $\frac{T_a - \beta}{\alpha - \beta}$ In honest EBL the minimum probability with which a particular schema is produced from a problem is p_{min} Thus the probability that SC is produced by EBL on a random problem of the appropriate type is at least $\frac{p_{min}(T_a - \beta)}{\alpha - \beta}$ QED

Lemma 5 plays an important role in the tractability result Informally, it limits how successfully an adequate schema can hide from random probing The higher a schema s adequacy the easier it will be to find This may at first seem surprising but it is in fact quite intuitive A schema of high adequacy must correctly apply to many problems and from lemma 4 each of these problems can give rise to the schema through EBL Thus random probing affords more opportunities to learn a highly adequate schema than one that is only marginally adequate

3 4 Single Schema Planning

We now wish to consider a very simple conformable schema-based planner, called a *single schema planner* It is given problems drawn from just one problem class and it may acquire at most one schema from its bias set by applying honest EBL to problems drawn from the world's distribution

Definition *single schema permissive planning* is restricted permissive planning in which the planner is a single schema planner

Theorem 2 Single schema permissive planning is tractable

By "tractable we mean that the sample complexity (the number of domain problems consumed by the algorithm) is independent of the cardinality of the (potentially very large) bias space and is polynomial!y bounded in other relevant parameters (T_a α β γ p, p_{min} $1/\varepsilon$ $1/\delta$) From the definition of restricted permissive planning single schema permissive planning must with high probability, hall with a correct answer

Proof We will show that the following algorithm performs single schema permissive planning and that it does so tractably as defined above

Algorithm SSPP $(\varepsilon, \delta, T_a)$

GLOBAL $\alpha\ \beta\ \gamma\ \rho\ p_{min}$

% Adequacy constants and expected EBL characteristics %

$$\text{Set } M = \max\left(3\left\lceil \frac{\alpha-\beta}{p_{min}(T_a-\beta)} \ln\left(\frac{6}{\delta}\right)\right\rceil\right),$$

$$K = \left\lceil \frac{\delta\rho M + 3 + \sqrt{6\delta\rho M + 9}}{\delta\rho^2} \right\rceil,$$

$$N = \left\lceil M \ \frac{2(\alpha-\gamma)}{\varepsilon^2 \delta} \right\rceil$$

% See text for M N K descriptions %

Set Tried=0 Set Schemas=0

WHILE (Tried<K) and (Schemas<M) DO

Set Tried=Tried+1

Get a problem P_h

% a problem for hypothesis generation %

Set SC SEBL(P_h) Set A_{SC}=0

IF SC≠Λ BEGIN

Set Schemas=Schemas+1

REPEAT N times

Get a problem P_t

% a problem for hypothesis testing %

Apply SC to P_t to yield PLAN

IF PLAN = Λ THEN

$A_{SC}=A_{SC}+\beta$

ELSE IF Execute(PLAN) achieves Goal(P_t) Then

$A_{SC}=A_{SC}+\alpha$

ELSE $A_{SC}=A_{SC}+\gamma$

END REPEAT

END BEGIN,IF

$A_{SC}=A_{SC}/N$

IF $A_{SC}>T_a$ THEN Exit SSPP with SC

END WHILE

Exit SSPP with Λ

Informally, the algorithm attempts to construct and test as many as M schemata M is chosen so that if an adequate schema exists it is unlikely to be missed in all M trials Each test employs N sample problems N is chosen so that the measured adequacy is unlikely to differ very much from the schema s true adequacy

Proof The algorithm always halts Recall the requirements $T_a>\beta$, $\varepsilon>0$ $\delta>0$ and $0<\rho<1$ The outer WHILE loop is repeated at most K times K is finite The embedded DO loop has the finite iteration target N The only conditional flow of control exits the outer loop prematurely Thus the algorithms time complexity is O(KN) As many as K problems are drawn in the outer loop For as many as M (with $M\leq K$) iterations the inner loop is executed N times drawing one problem in each iteration Thus the algorithm halls consuming at most K+MN domain problems The parameters M K and N are assigned values which are independent of the cardinality of the bias space and polynomial in the other relevant parameters

It remains to show that the probability of producing an incorrect answer is bounded by 6 There are two ways the algorithm might yield an incorrect answer 1) It may output Λ failing to Find an adequate schema when one exists or 2) it may halt with an inadequate schema (whether or not an adequate one exists) The first error is a false negative call it Ffn The second error Ffp is a false positive We must show that *Pr(F,) + Pr(F,)<,S* It suffices to show $\Pr(F_{fn}) \leq \delta/2$ and $\Pr(F_{fp}) \leq \delta/2$

We first consider Ff_p This results if any one of the M hypothesized schemata is incorrectly judged to be adequate To avoid a false positive it suffices that $\sum_M \delta_i \leq \delta/2$ where each d_t is the probability that the *1'th* hypothesized schema will be falsely confirmed as adequate Thus it suffices to show that $\delta_i \leq \delta/2M$ for each i (4)

The algorithm estimates the adequacy of each hypothesized schema by applying it to N sample problems The *8,* s represent the probability that the result of the *i th* test falsely confirmed by the observed data Statistics assures us that a sufficiently large N will provide the required accuracy We show thai such an N respects our conditions of tractability

Provided the variance over a randomly sampled distribution is finite Chebyshev s inequality provides a useful relation for PAC-style proofs (Natarajan 1991) Chebyshev s inequality sidles thai $\Pr\left(\left|\mu-\mu_{\bar{X}}\right| \geq \Delta\right) \leq \frac{\sigma^2}{\Delta^2 N}$ where μ and σ^2 are the true mean and true variance respectively of a (not necessarily normal) random variable X The symbol $\mu_{\bar{X}}$ represents the measured mean of N samples of X Δ is positive The relation quantifies the probability of observing a measured mean very differeni from the true mean

Let SC be a hypothesized Schema under evaluation A_s be the computed sample adequacy of SC over N random problems and At be the (unknown) true adequacy of SC over the underlying distribution The underlying distribution of adequacies is a trinomial of finite values so o"2 is also finite Chebyshev s inequality applies For any positive A

$$\Pr\left(\left|A_S - A_t\right| \geq \Delta\right) \leq \frac{\alpha-\gamma}{\Delta^2 N} \qquad (5)$$

A necessary and sufficient condition for SC to be judged a false positive is that its sample adequacy is measured to be at least the desired adequacy threshold but its true adequacy is less than the threshold by at least E

$(A_S \geq T_a) \wedge (A_t < T_a - \varepsilon)$ which implifies to

$$A_S > A_t + \varepsilon$$

Thus a sufficient condition for avoiding a false positive is that for each test

$$\left|A_S - A_t\right| < \varepsilon \qquad (6)$$

Inequality (5) provides a bound on the probability that inequality (6) is violated $\Pr\left(\left|A_S - A_t\right| \geq \varepsilon\right) \leq \frac{\alpha - \gamma}{\varepsilon^2 N}$ For N to be sufficient we require this probability to satisfy Inequality (4) $\frac{\alpha-\gamma}{\varepsilon^2 N} \leq \frac{\delta}{2M}$ Thus N must be chosen so $N \geq M\ \frac{2(\alpha-\gamma)}{\varepsilon^2\delta}$ which is respected by the algorithm Note that N is independent of the cardinality of the bias space (the number of entertainable schemata) and that it is polynomially bounded provided M is

Next we turn to M and the probability of a false negative F_{fn} For a false negative condition the algorithm incorrectly claims that no adequate schema exists There are three ways the algorithm might result in a false negative First the K sampled problems may be sufficiently difficult that fewer than M can be solved In this case the algorithm cannot construct the requisite M random hypothesis schemata Call this false negative failure F_{fn1} The second type of failure F_{fn2}, occurs when no adequate schema is entertained among the M hypotheses Finally one or more adequate schemata may be hypothesized but are judged to be inadequate by the statistical adequacy tests F_{fn3} We require that $\sum_{i=1}^{3}\Pr(F_{fni}) \leq \delta/2$ It suffices to show $\Pr(F_{fni}) \leq \delta/6$ for each i

First we consider F_{fn1} We will show for the given method of selecting K that with probability $1-\delta/6$ at least M of the K sampled problems are solvable We assume the difficulty of a problem is an independent property of the problem Provided the resources, R allocated to the EBL planner are roughly adequate for the intrinsic difficulties of the problem class, the assumption seems reasonable

Let us define a new random variable S S_i is 1 if problem i is solved within the resources allowed and 0 if it is not The mean of S of course is ρ, the expected proportion of problems that can be solved We must limit the chance of getting fewer than M solvable problems to $\frac{\delta}{6}$ We require

$$\Pr\left(\rho_{\bar{S}} \leq \frac{M}{K}\right) \leq \frac{\delta}{6} \qquad \text{or equivalently}$$

$$\Pr\left(\rho - \rho_{\bar{S}} \geq \frac{K\rho - M}{K}\right) \leq \frac{\delta}{6} \qquad (7)$$

Chebyshev applies to (7) if we let the number of samples be K define $\Delta = \frac{K\rho - M}{K}$ and note that the variance of S is at most 1 Chebyshev assures us that

$$\Pr\left(\left|\rho - \rho_{\bar{S}}\right| \geq \frac{K\rho - M}{K}\right) \leq \frac{1}{\left(\frac{K\rho - M}{K}\right)^2 K} \qquad (8)$$

The left hand side of (7) is weaker than (8) since Chebyshev also provides a bound on how much the observed solution rate $\rho_{\bar{S}}$ can be larger than the true rate ρ We require only that the observed rate not often be much smaller Thus we successfully limit F_{fn1} if K is chosen to satisfy $\frac{1}{\left(\frac{K\rho - M}{K}\right)^2 K} \leq \frac{\delta}{6}$ This simplifies to

$$K \geq \frac{\delta\rho M + 3 + \sqrt{6\delta\rho M + 9}}{\delta\rho^2}$$

The SSPP algorithm chooses K consistent with this bound which is independent of the size of the bias space and polynomial in the other parameters M again appears linearly so a tractable number of examples can probabilistically guarantee a low probability of F_{fn1} only if M is also tractable

We require F_{fn2} the possibility that an adequate schema is not among the M hypothesized schemata to be unlikely $\Pr(F_{fn2}) \leq \delta/6$ Let SC be an adequate schema We know an adequate schema exists or the false negative condition could not arise By lemma 5 if a random schema is acquired using a randomly sampled problem the probability that the constructed schema will be the adequate schema SC is at least $\frac{p_{\min}(T_a - \beta)}{\alpha - \beta}$ Conversely the probability that the randomly acquired schema is not SC is no greater than $1 - \frac{p_{\min}(T_a - \beta)}{\alpha - \beta}$ Repeated randomly acquired schemata are independent events Thus the probability that all M attempts fail to acquire SC is no greater than $\left(1 - \frac{p_{\min}(T_a - \beta)}{\alpha - \beta}\right)^M$ It suffices to choose M to satisfy

$$\left(1 - \frac{p_{\min}(T_a - \beta)}{\alpha - \beta}\right)^M \leq \frac{\delta}{6} \qquad \text{or sufficiently}$$

$$M \geq \frac{\alpha - \beta}{p_{\min}(T_a - \beta)} \ln\left(\frac{6}{\delta}\right)$$ Thus M is also tractable

Finally there is the possibility of F_{fn3} that an entertained adequate schema fails its confirmation test and is incorrectly judged to be inadequate We require

$Pr(F_{fn3}) \leq \delta/6$ Recall again that T_a is the desired adequacy threshold and that e is E-adequacy parameter A necessary condition for F_{fn3} failure is that a schema s sample adequacy is measured to be less than the desired adequacy threshold but its true adequacy is at least E greater than the threshold $(A_s < T_a) \wedge (A_t \geq T_a + \varepsilon)$ or simplifying, $A_s < A_t - \varepsilon$ A sufficient condition for avoiding this form of false positive is Inequality (6) But this is precisely the false positive error condition There we required a sufficiently large N to bound the error probability of each test by $\delta/2M$ Here we need to bound the error probability of each test by $\delta/6$ It we require M to be at least 3, the error probabilities tor both the false positive condition a n F_{fn3} I I be bounded as desired M is chosen to be at least 5

Thus the probability for a false negative and the probability for a false positive are each required to be no greater than $\delta/2$ so the algorithm produces the correct answer with probability at least $1-\delta$ QED

4 Conclusions

Much previous research has applied machine learning techniques to planning Usually ML is employed to improve the accuracy of a planners representations (e g (Gil 1994)) This alters the planners projection ability rather than its bias The chunking of control knowledge (e g SOAR (Laird *el* al 1987) and PRODIGY/EBL (Minton 1988)) can be viewed as altering planner bias as can some case-based approaches (Allen and Langley 1990 Hammond e(al 1990) though in these systems there is no formal connection to planner adequacy which is central to permissive planning

In addition lo our analytic results we have implemented a real-world single-schema permissive planning system GRASPER a robot system that learns lo pick up novel objects Its model of world change is simple and imperfect A bias is acquired that prefers to select among other things grasp points nearer the center of mass grasp forces higher than believed necessary and wider than needed gripper openings on approach to the target piece

While single schema permissive planning may seem narrow the single schema results are immediately extendible to a schema planner with one schema per problem class With little difficulty the results also extend to a schema planner with any fixed maximum number of schemata for each problem type

The permissive planning approach offers to bring plan execution and real-world behavior within the framework of classical planning We show that one natural way to realize permissive planning is with the combination of EBL and a schema planner The tractability proof guarantees that permissive planning scales - that the benefits are not the manifestation of a simple domain

Acknowledgement

The authors wish to thank Lenny Pitt for helpful discussions and insights and Renee Baillargeon for comments on the manuscript

References

Allen J and P Langley (1990) Integrating Memory and Search in Planning Workshop on Innovative Approaches to Planning Scheduling and Control San Diego CA DARPA

Bechhofer R (1954) A Single Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances *Annals of Mathematical Statistics* 25(1) 16-39

Bennett S and G F DeJong(199l) Comparing Stochastic Planning to the Acquisition of Increasingly Permissive Plans for Complex Uncertain Domains Eighth International Workshop on Machine Learning Evansion IL

Bennett S W (1993) Permissive Planning A Machine Learning Approach to Planning in Complex Real World Domains Ph D Thesis University of Illinois Electrical &. Computer Engineering Dept

DeJong G F and S Bennett (1993) Permissive Planning A Machine Learning Approach to Linking Internal and External Worlds Eleventh National Conference on Artificial Intelligence Washington D C MTT Press

Gil Y (1994) Learning by Experimentation Incremental Refinement of Incomplete Planning Domains Eleventh International Conference on Machine Learning Rutgers Universitv Morgan Kaufmann

Hammond K el al (1990) Integrating Planning and Acting in a Case Based Framework Eighth National Conference on Artificial Intelligence pp 292-297

Laird J et al (1987) SOAR An Architecture for General Intelligence *Artificial Intelligence* 33(1) 1-64

Minton S (1988) *Learning Search Control Knowledge An Explanation Based Approach* Norwell MA Kluwer Academic Publishers

Natarajan B (1991) *Machine Learning a Theoretical Approach* San Mateo Morgan Kaufmann

Presman E L and 1 N Sonin(1990) *Sequential Control with Incomplete Information* New York Academic Press

Valiant L (1984) A Theory of the Learnable *Communications of the ACM 27*(11) 1134-1 [42