

Metareasoning for Concurrent Planning and Execution

Dylan O’Ceallaigh and Wheeler Ruml



UNIVERSITY *of* NEW HAMPSHIRE

Department of Computer Science

Grateful thanks to NSF and LAAS-CNRS for support.

Concurrent Planning and Acting

When to plan and when to act?

Introduction

Planning & Acting

Problem Setting

When to Plan More?

How Many Actions?

Conclusion

Concurrent Planning and Acting

When to plan and when to act?

Assumption: always planning (dedicated core)

[Introduction](#)

[Planning & Acting](#)

[Problem Setting](#)

[When to Plan More?](#)

[How Many Actions?](#)

[Conclusion](#)

Concurrent Planning and Acting

Introduction

■ Planning & Acting

■ Problem Setting

When to Plan More?

How Many Actions?

Conclusion

When to plan and when to act?

Assumption: always planning (dedicated core)

Question: **when to commit?**

Concurrent Planning and Acting

Introduction

■ Planning & Acting

■ Problem Setting

When to Plan More?

How Many Actions?

Conclusion

When to plan and when to act?

Assumption: always planning (dedicated core)

Question: **when to commit?**

1. off-line: complete plan before acting
implicit *identity action* that preserves state

Concurrent Planning and Acting

Introduction

■ Planning & Acting

■ Problem Setting

When to Plan More?

How Many Actions?

Conclusion

When to plan and when to act?

Assumption: always planning (dedicated core)

Question: **when to commit?**

1. off-line: complete plan before acting
implicit *identity action* that preserves state
2. real-time: plan incrementally, commit to current best
never execute *identity*, plan while acting

Concurrent Planning and Acting

Introduction

■ Planning & Acting

■ Problem Setting

When to Plan More?

How Many Actions?

Conclusion

When to plan and when to act?

Assumption: always planning (dedicated core)

Question: **when to commit?**

1. off-line: complete plan before acting
implicit *identity action* that preserves state
2. real-time: plan incrementally, commit to current best
never execute *identity*, plan while acting

How to choose?

Is there a middle ground?

Can we plan/commit dynamically?

How can we decide in a principled way?

Concurrent Planning and Acting

Introduction

■ Planning & Acting

■ Problem Setting

When to Plan More?

How Many Actions?

Conclusion

When to plan and when to act?

Assumption: always planning (dedicated core)

Question: **when to commit?**

1. off-line: complete plan before acting
implicit *identity action* that preserves state
2. real-time: plan incrementally, commit to current best
never execute *identity*, plan while acting

How to choose?

Is there a middle ground?

Can we plan/commit dynamically?

How can we decide in a principled way?

Metareasoning!

The Problem Setting

Introduction

■ Planning & Acting

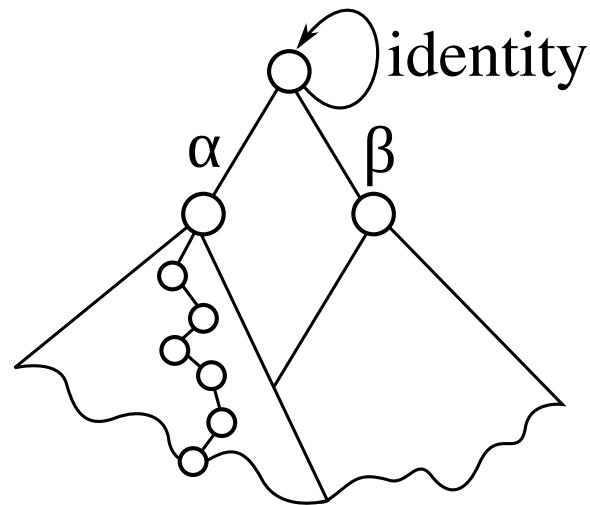
■ Problem Setting

When to Plan More?

How Many Actions?

Conclusion

1. planning as forward state-space heuristic search
2. minimize goal achievement time (GAT)
action 'cost' = duration
3. access to an inadmissible heuristic \hat{h} ($+ g = \hat{f}$)
4. for simplicity: known deterministic world, serial plan



The Problem Setting

Introduction

■ Planning & Acting

■ Problem Setting

When to Plan More?

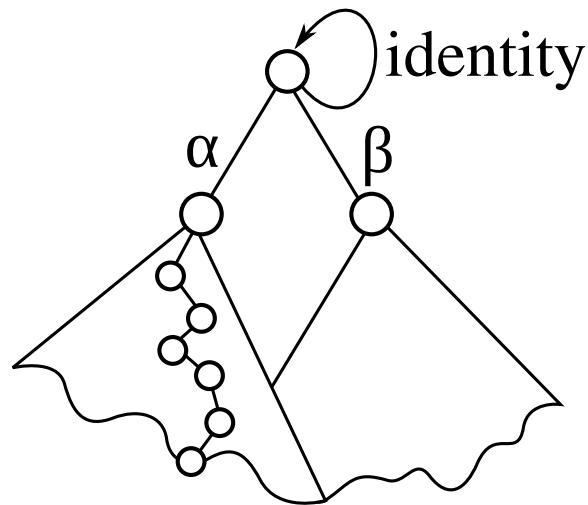
How Many Actions?

Conclusion

1. planning as forward state-space heuristic search
2. minimize goal achievement time (GAT)
action 'cost' = duration
3. access to an inadmissible heuristic \hat{h} ($+ g = \hat{f}$)
4. for simplicity: known deterministic world, serial plan

Central acting decisions:

1. execute (a) current-best action or (b) identity action?
2. if (a), how many actions?



Metareasoning (Russell and Wefald, 1991)

plan when it appears worthwhile!

Introduction

When to Plan More?

■ Metareasoning

■ Estimating Belief

■ IMR

■ Simple Problems

■ Large Problems 1

■ Large Problems 2

How Many Actions?

Conclusion

Metareasoning (Russell and Wefald, 1991)

plan when it appears worthwhile!

plan when **expected GAT reduction** > planning time

Introduction

When to Plan More?

■ Metareasoning

■ Estimating Belief

■ IMR

■ Simple Problems

■ Large Problems 1

■ Large Problems 2

How Many Actions?

Conclusion

Metareasoning (Russell and Wefald, 1991)

plan when it appears worthwhile!

plan when **expected GAT reduction** > planning time

GAT reduction depends on whether $\hat{f}(\alpha) > \hat{f}(\beta)$ after search
and if so, $\hat{f}(\alpha) - \hat{f}(\beta)$

Introduction

When to Plan More?

■ Metareasoning

■ Estimating Belief

■ IMR

■ Simple Problems

■ Large Problems 1

■ Large Problems 2

How Many Actions?

Conclusion

Metareasoning (Russell and Wefald, 1991)

plan when it appears worthwhile!

plan when **expected GAT reduction** > planning time

GAT reduction depends on whether $\hat{f}(\alpha) > \hat{f}(\beta)$ after search
and if so, $\hat{f}(\alpha) - \hat{f}(\beta)$

More precisely, if x_α, x_β are possible \hat{f} values after search:

$$b(x_\alpha, x_\beta) = \begin{cases} 0 & \text{if } x_\alpha \leq x_\beta \\ x_\alpha - x_\beta & \text{otherwise} \end{cases}$$

If $P_{\hat{f}(n)}$ represents belief over future value,

$$B = \int_{x_\alpha} P_{\hat{f}(\alpha)}(x_\alpha) \int_{x_\beta} P_{\hat{f}(\beta)}(x_\beta) b(x_\alpha, x_\beta) dx_\beta dx_\alpha$$

Introduction

When to Plan More?

■ Metareasoning

■ Estimating Belief

■ IMR

■ Simple Problems

■ Large Problems 1

■ Large Problems 2

How Many Actions?

Conclusion

Estimating Belief Distributions

how to estimate where $\hat{f}(n)$ will be after search?

[Introduction](#)

[When to Plan More?](#)

■ Metareasoning

■ **Estimating Belief**

■ IMR

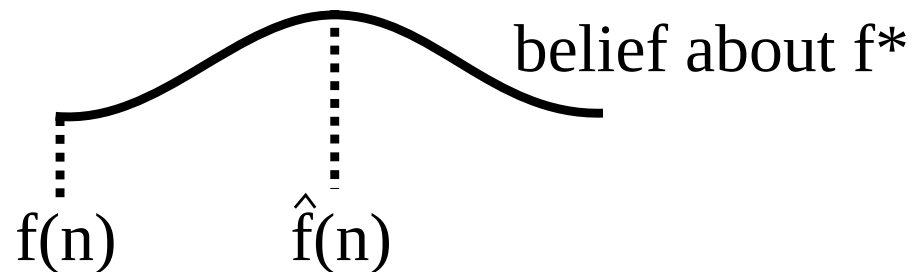
■ Simple Problems

■ Large Problems 1

■ Large Problems 2

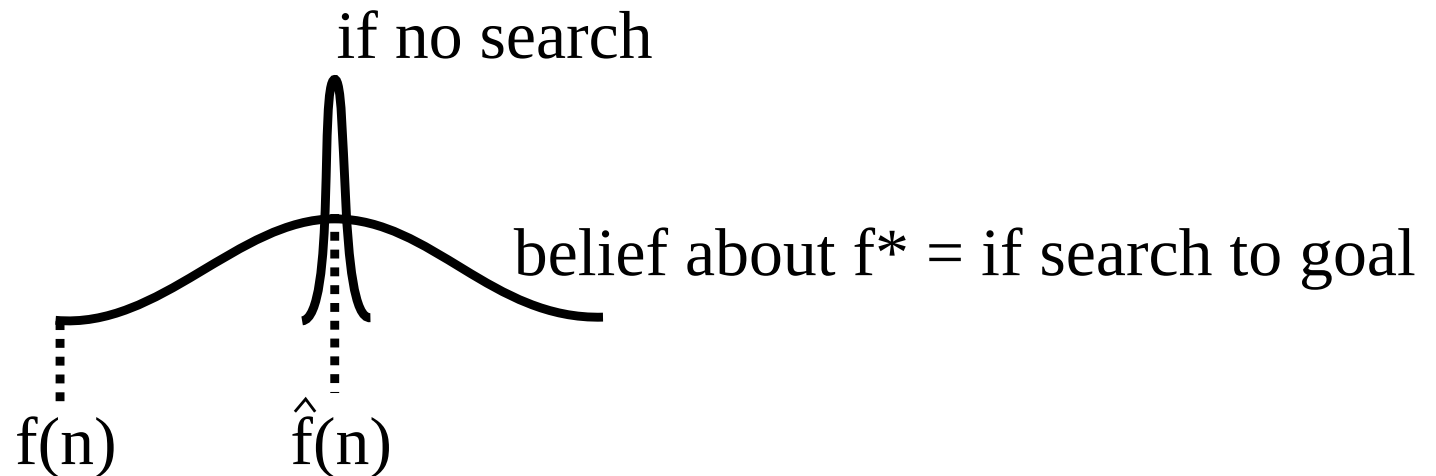
[How Many Actions?](#)

[Conclusion](#)



Estimating Belief Distributions

how to estimate where $\hat{f}(n)$ will be after search?



[Introduction](#)

[When to Plan More?](#)

■ Metareasoning

■ **Estimating Belief**

■ IMR

■ Simple Problems

■ Large Problems 1

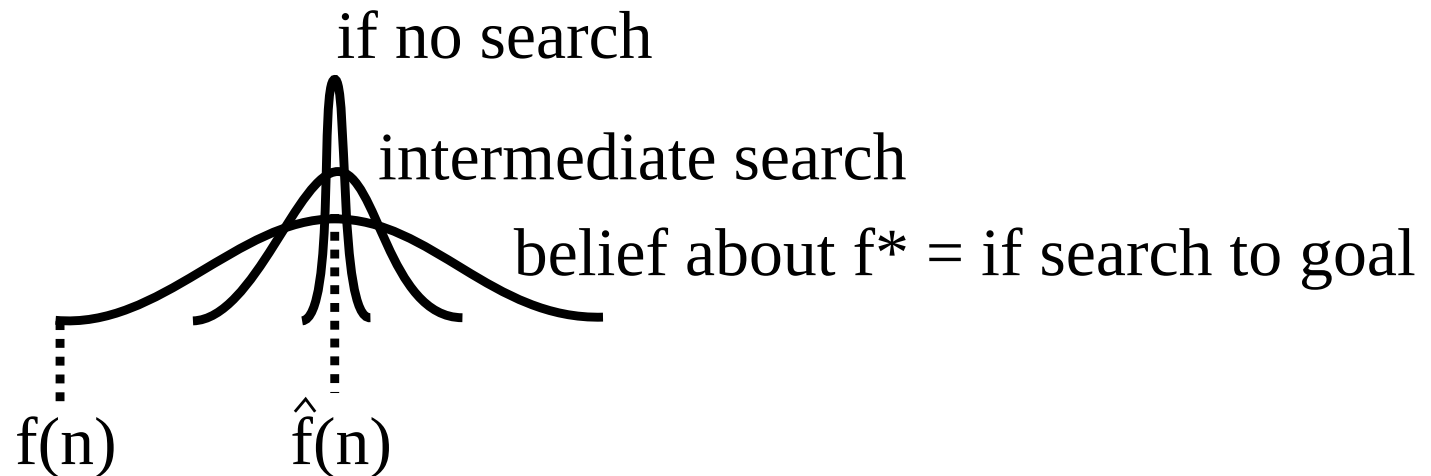
■ Large Problems 2

[How Many Actions?](#)

[Conclusion](#)

Estimating Belief Distributions

how to estimate where $\hat{f}(n)$ will be after search?



[see paper for details]

[Introduction](#)

[When to Plan More?](#)

■ Metareasoning

■ **Estimating Belief**

■ IMR

■ Simple Problems

■ Large Problems 1

■ Large Problems 2

[How Many Actions?](#)

[Conclusion](#)

Sketch of IMR (Identity MetaReasoning)

Introduction

When to Plan More?

■ Metareasoning

■ Estimating Belief

■ IMR

■ Simple Problems

■ Large Problems 1

■ Large Problems 2

How Many Actions?

Conclusion

variant of Dynamic \hat{f} real-time search (Burns et al, SoCS-13)

1. until a goal is reached
2. best-first search on \hat{f} until *time bound*
3. if *identity* is applicable and $B > t_{identity}$
4. $a \leftarrow identity$
5. else
6. $a \leftarrow$ first action in best partial plan
7. update heuristic values
8. reset search
9. *time bound* $\leftarrow a$'s duration
10. start executing a

Results on Simple Benchmarks

Introduction

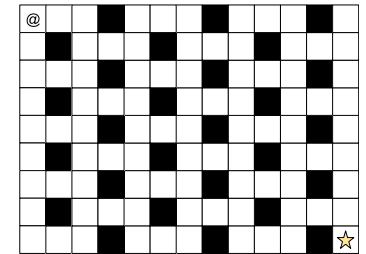
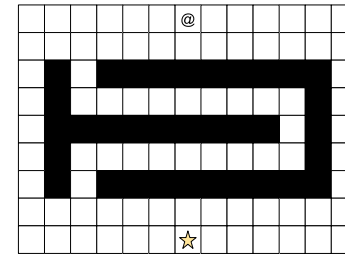
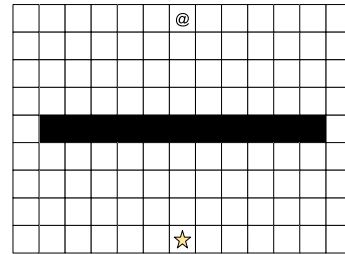
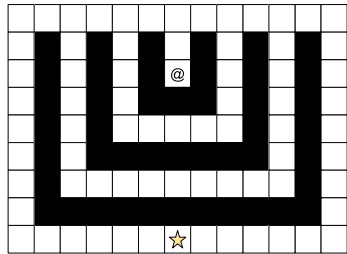
When to Plan More?

- Metareasoning
- Estimating Belief
- IMR
- Simple Problems
- Large Problems 1
- Large Problems 2

How Many Actions?

Conclusion

Sketches:

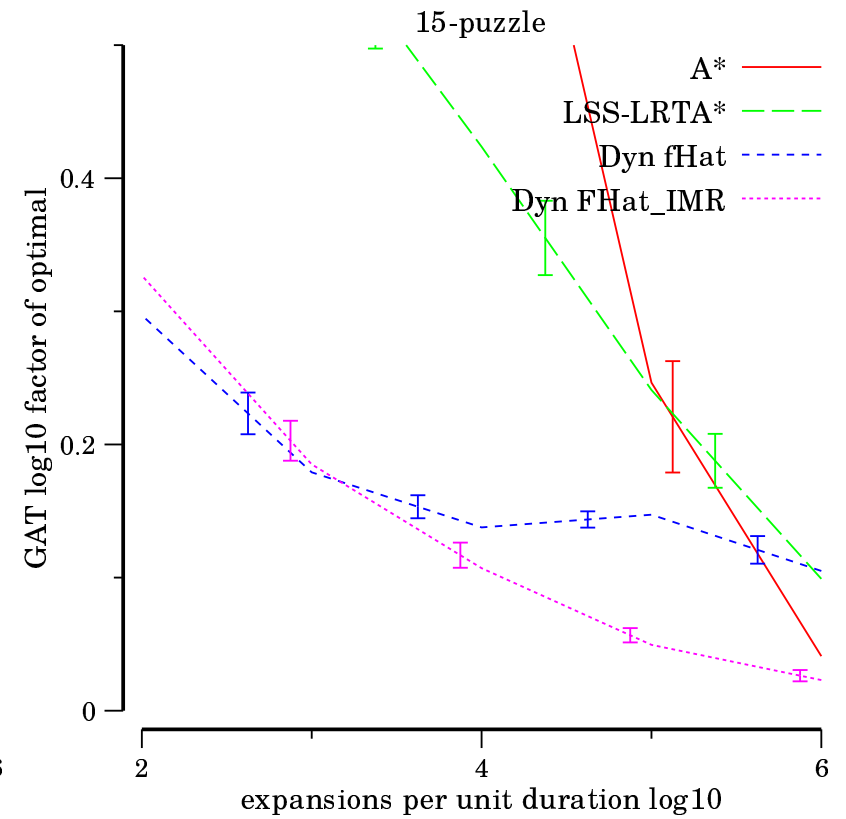
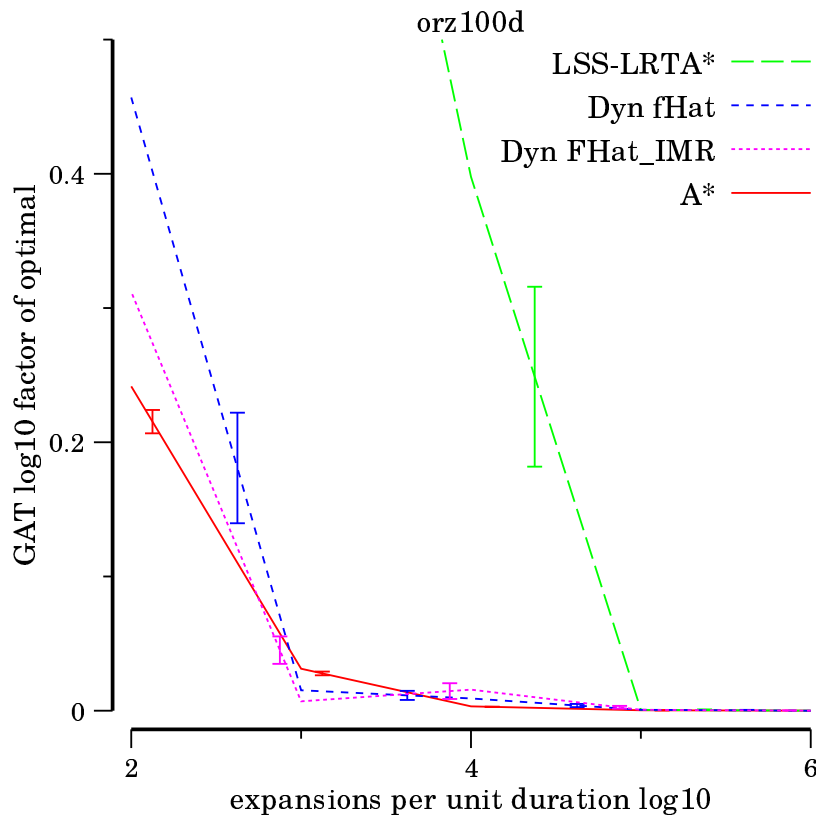


	A*	real-time		IMR
		LSS-LRTA*	\hat{f}	
cups	166	3,500	5,322	970
wall	102	523	717	101
slalom	177	382	638	161
uniform	29,578	3,195	2,997	2,997

IMR adapts from off-line to real-time!

Results on Larger Benchmarks (1/2)

- Introduction
- When to Plan More?
 - Metareasoning
 - Estimating Belief
 - IMR
 - Simple Problems
 - Large Problems 1
 - Large Problems 2
- How Many Actions?
- Conclusion



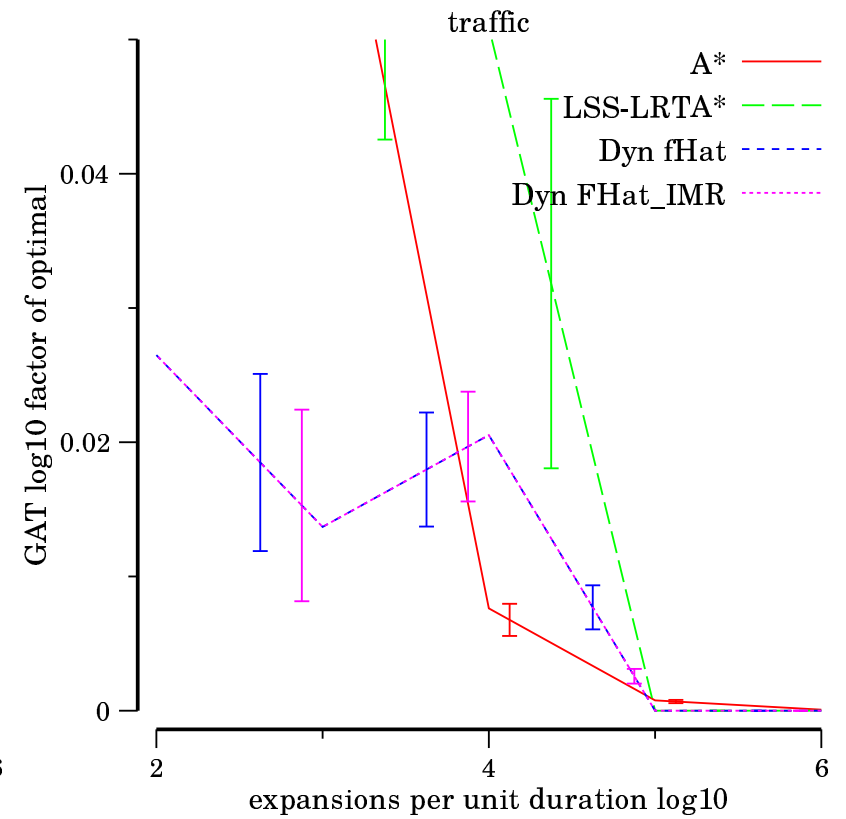
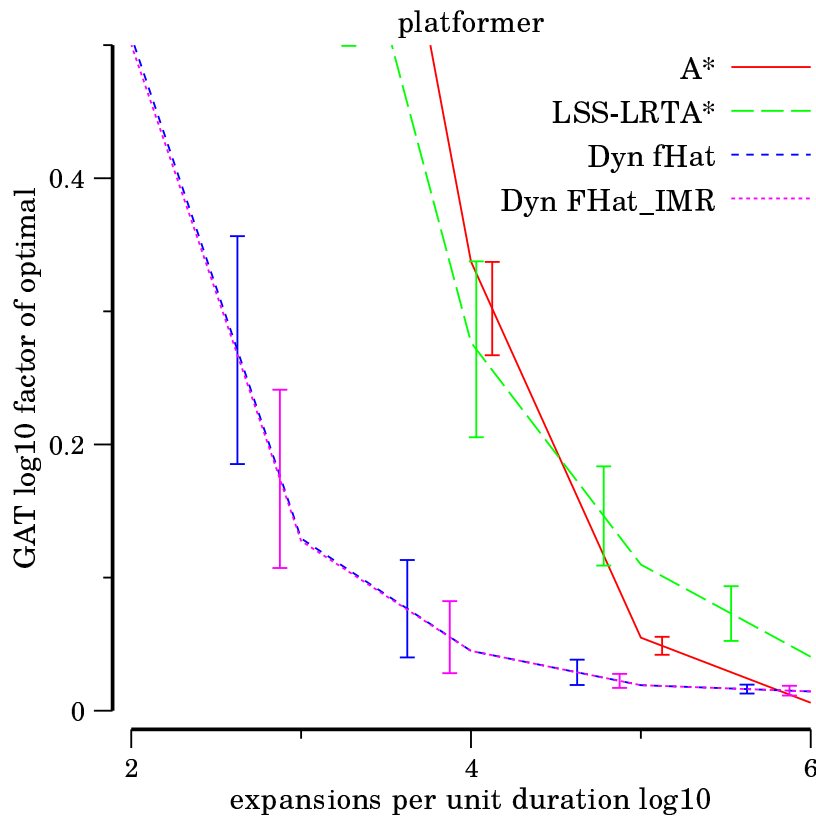
IMR approaches A*!

	1	2	3
4	5	6	7
8	9	10	11
12	13	14	15

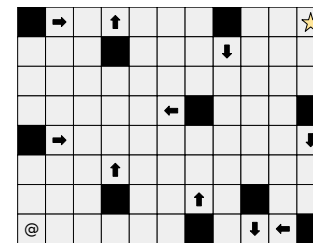
IMR adapts!

Results on Larger Benchmarks (2/2)

- Introduction
- When to Plan More?
 - Metareasoning
 - Estimating Belief
 - IMR
 - Simple Problems
 - Large Problems 1
 - Large Problems 2
- How Many Actions?
- Conclusion



IMR mimics real-time



no identity actions

How Many Actions to Commit To?

[Introduction](#)

[When to Plan More?](#)

[How Many Actions?](#)

■ **How Many?**

■ Simple Problems

■ Large Problems 1

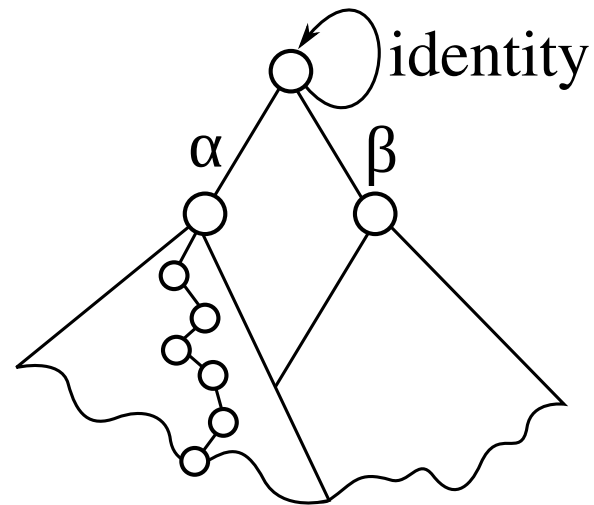
■ Large Problems 2

[Conclusion](#)

Consider each node along partial path

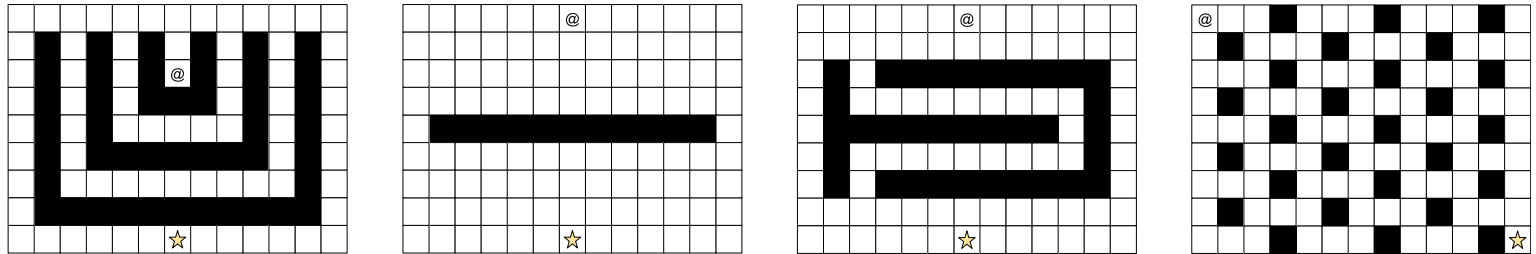
Stop at the first where planning is preferred

When combined with previous method: **Mo'RTS**



Results on Simple Benchmarks

Sketches:



	A*	real-time			IMR	Mo'RTS
		LSS-LRTA*	\hat{f}			
cups	166	3,500	5,322	970	241	
wall	102	523	717	101	140	
slalom	177	382	638	161	161	
uniform	29,578	3,195	2,997	2,997	2,997	

Mo'RTS perhaps improves slightly over IMR

Introduction

When to Plan More?

How Many Actions?

■ How Many?

■ Simple Problems

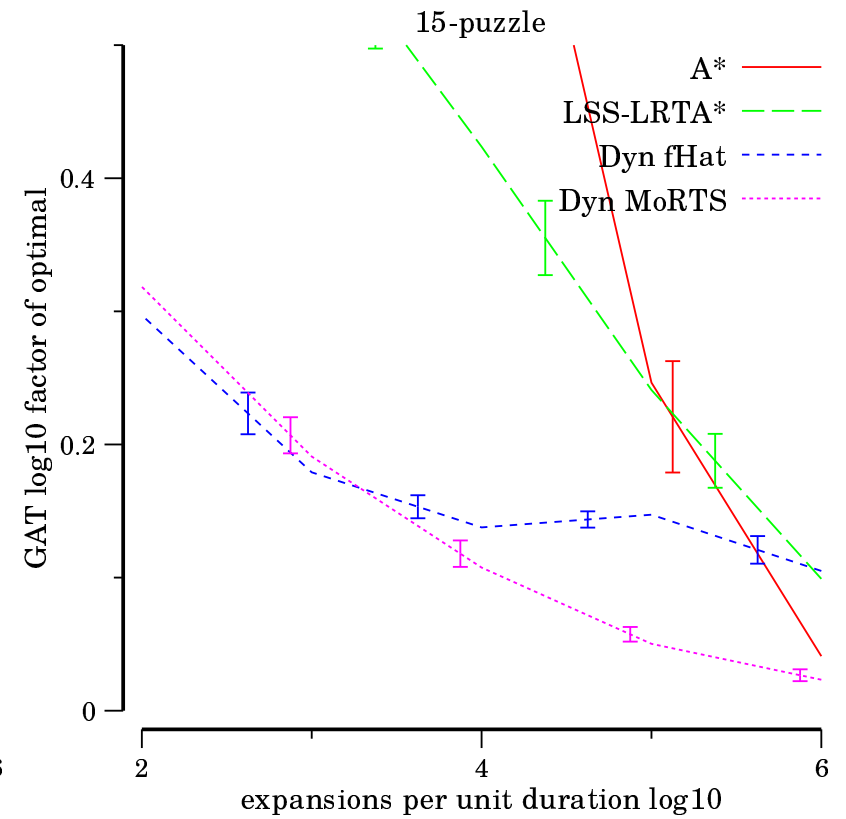
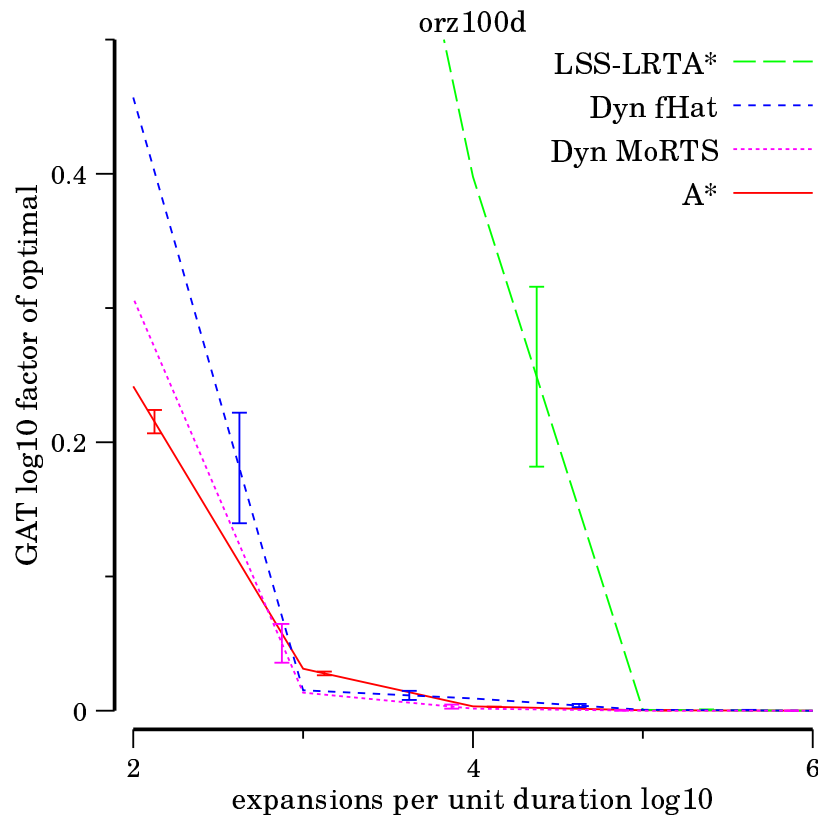
■ Large Problems 1

■ Large Problems 2

Conclusion

Results on Larger Benchmarks(1/2)

- [Introduction](#)
- [When to Plan More?](#)
- [How Many Actions?](#)
- How Many?
 - Simple Problems
 - **Large Problems 1**
 - Large Problems 2
- [Conclusion](#)



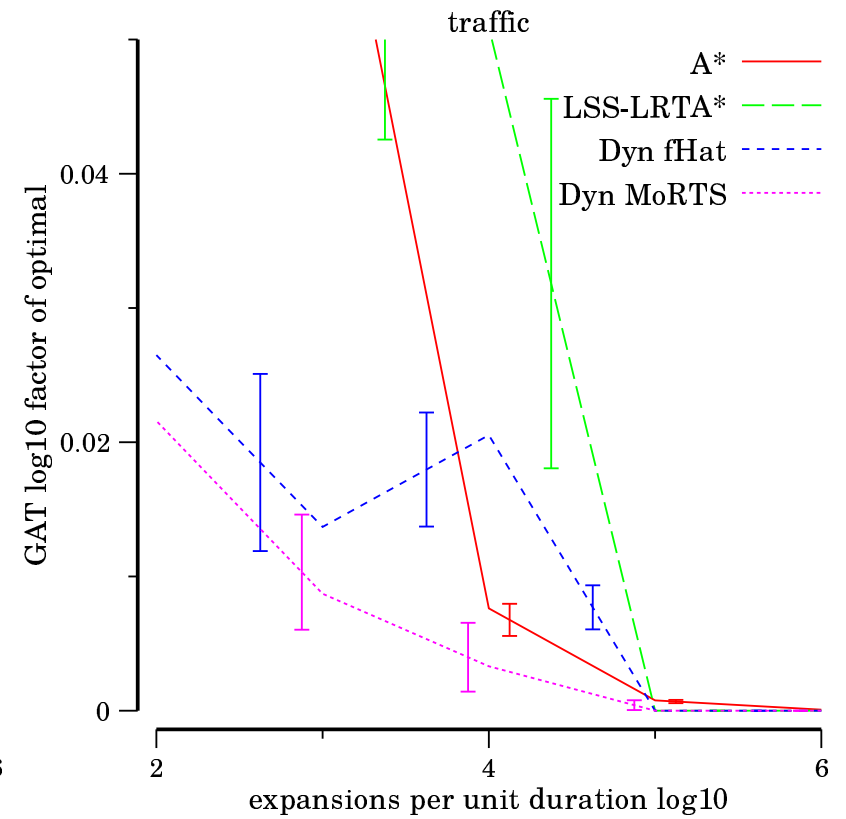
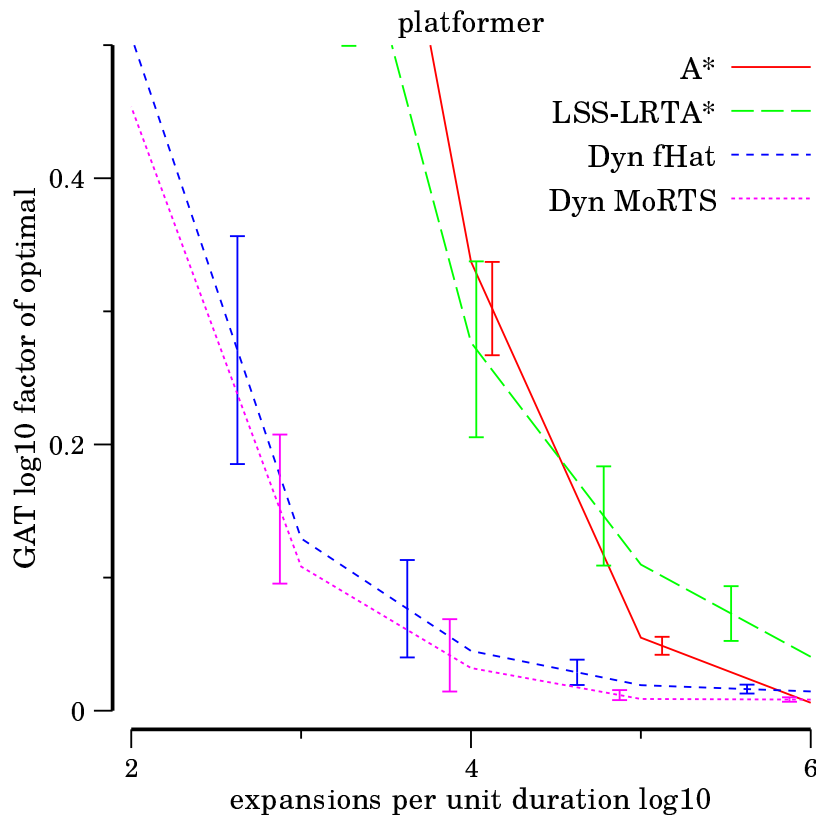
Mo'RTS
matches IMR

	1	2	3
4	5	6	7
8	9	10	11
12	13	14	15

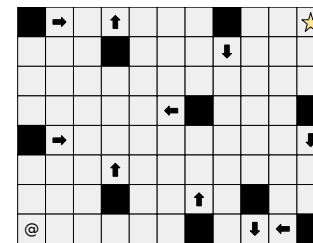
Mo'RTS
matches IMR

Results on Larger Benchmarks (2/2)

- [Introduction](#)
- [When to Plan More?](#)
- [How Many Actions?](#)
- How Many?
- Simple Problems
- Large Problems 1
- **Large Problems 2**
- [Conclusion](#)



possible slight improvement



Mo'RTS much better

Conclusions

[Introduction](#)

[When to Plan More?](#)

[How Many Actions?](#)

[Conclusion](#)

■ [Conclusions](#)

Objective: Minimize time to goal achievement

1. plan then act: **Bugsy** (Burns, Ruml, and Do, JAIR 2013)
2. concurrent planning and acting: **Mo'RTS** (this work)

Approach: Metareasoning

1. beautiful principle
2. provides state-of-the-art results in practice
3. should be integrated into the planner

Possible extensions

1. non-deterministic and partially-known settings
2. 'not-quite-identity' actions
3. plan-space planning

Practical metareasoning for adaptive deliberation!

Introduction

When to Plan More?

How Many Actions?

Conclusion

Extra Slides

■ Assumptions

■ DTA*

Extra Slides

Assumptions

[Introduction](#)

[When to Plan More?](#)

[How Many Actions?](#)

[Conclusion](#)

[Extra Slides](#)

■ [Assumptions](#)

■ [DTA*](#)

- inadmissible \hat{h}
- Gaussian belief
- linear variance reduction with lookahead
- estimate of future expansion delay
- cost of committing before frontier
- identity and length of commitment are separate decisions
- only consider acting at action end times

Decision-Theoretic A* (Russell and Wefald, 1991)

[Introduction](#)

[When to Plan More?](#)

[How Many Actions?](#)

[Conclusion](#)

[Extra Slides](#)

■ Assumptions

■ DTA*

- same basic principle
- based on older RTA* instead of Dynamic \hat{f}
 - ◆ assumes disjoint subtrees beneath current actions
 - ◆ assumes admissible h
 - ◆ non-A* lookahead
- estimates effect of search using training data