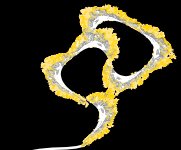
UNIVERSITY OF TWENTE.

Exploration and the Shape of Information

Inaugural Lecture

Prof.dr. W.M. Koolen

13 March 2025



Mathematical Machine Learning

Online Learning Bandits Pure Exploration Optimisation Zero-sum games Testing Martingales
Probability Algorithms

Mathematical Machine Learning

Online Learning Bandits Pure Exploration Optimisation Zero-sum games Testing Martingales Probability Algorithms

Two Questions

Q1: What do we know?

Q2: What should we **do**?

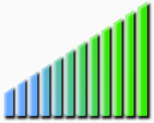
Q1: What do we know?

Q2: What should we **do**?

Philosophy $\Rightarrow$ Mathematics/Statistics $\Rightarrow$ Algorithms

Modern Machine Learning

Our Lives Will be Transformed



- Recent progress in Machine Learning is amazing.
- It has become **useful** ...
- in a spectrum of tasks: images, text, dialogue, programming, ...
- Fast and furious.

can you help me generate an image of a professor in dutch gown and cap on a red racing bicycle?

Here's the image of a professor in a Dutch academic gown and cap, riding a red racing bicycle. Let me know if you'd like any tweaks!

can you make the gown black with a green bevel, and can you make the bicycle more aero?

Here's the updated image with a black gown featuring a green bevel and a more aerodynamic red racing bicycle. Let me know if you'd like any further refinements!

Can you make the gown in the style of the University of Twente?

Here's the updated image with the gown in the style of the University of Twente. Let me know if you need any further refinements!

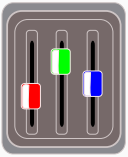
How Does that Work

Data:

$$(x_1, y_1), \ldots, (x_n, y_n) \qquad n \approx \text{internet}$$

Model:

$$Y = f(X; \theta_1, \ldots, \theta_m) \qquad m \approx \text{billion}$$



I am Excited

My own field will be the reason that **my job** will **transform significantly** in < 5 years.

I am Excited

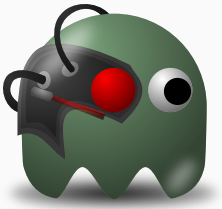
My own field will be the reason that **my job** will **transform significantly** in < 5 years.

And that is *exciting / natural / good / interesting*

I am Excited

My own field will be the reason that **my job** will **transform significantly** in < 5 years.

And that is *exciting* / *natural* / *good* / *interesting*



Challenges

Many scientific and societal questions

- Why does it work so **well**?
- When and how does it **fail**?
- Efficiency?
- Data ownership?
- Transparency? Accountability? Explainability? Fairness? Alignment?
- Who benefits?

Today's Question

After training on **all available data** . . . ?

Is there a way to improve further?

Today's Question

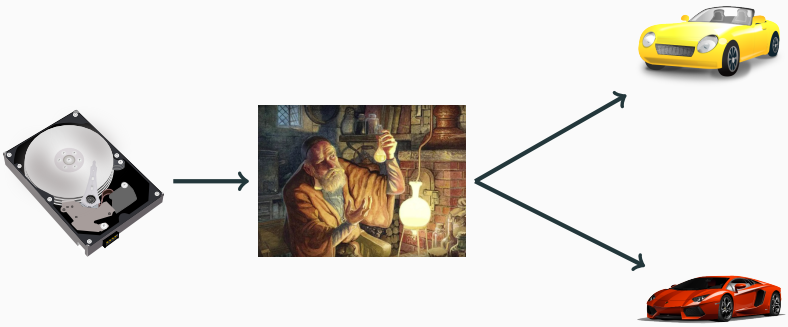
After training on **all available data** . . . ?

Is there a way to improve further?

Yes! Do experiments to generate **new** data!

Q1: What do we know?

Starting Point: Two Versions of Autonomous Driving Software



World Model

We assume each version has an **unknown** crash rate

/10.000

21.00

22.85

World Model

We assume each version has an **unknown** crash rate

/10.000

21.00

22.85

Together these determine the best version on average,  .

Learning the Best Version by Observing Driving Results

Learning the Best Version by Observing Driving Results

1  

Learning the Best Version by Observing Driving Results

1  

2  

Learning the Best Version by Observing Driving Results

1  

2  

3  

Learning the Best Version by Observing Driving Results

1  

2  

3  

4  

Learning the Best Version by Observing Driving Results

1  

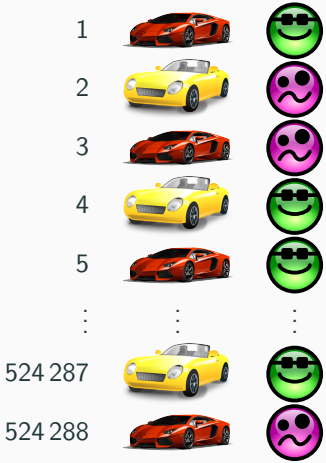
2  

3  

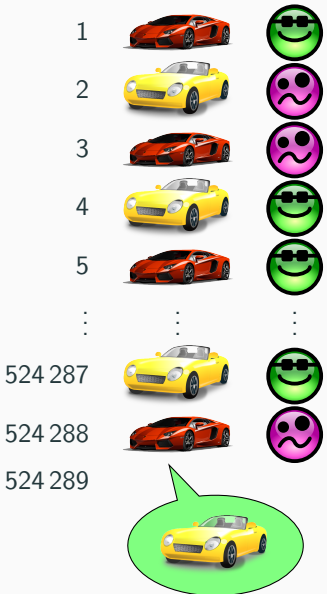
4  

5  

Learning the Best Version by Observing Driving Results

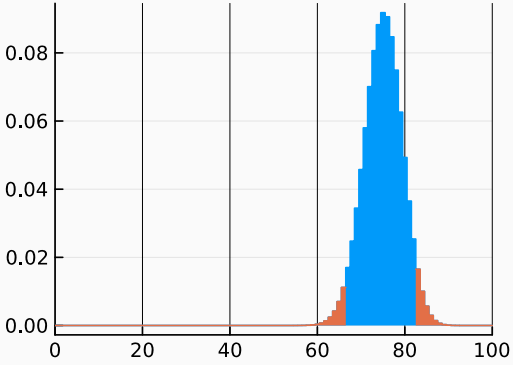


Learning the Best Version by Observing Driving Results



Frequentist Confidence Intervals

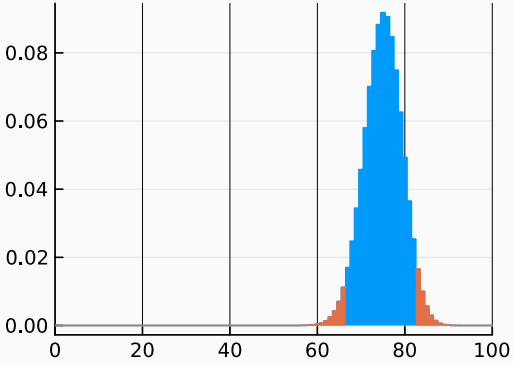
At any point, what do we know?



- Say the success rate is 3/4 and we sample 100 times
- then we will **typically** see a number of successes in the blue centre (67 – 82)
- while a number of successes in either red tail is **weird**

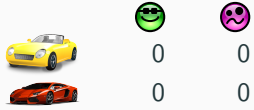
Frequentist Confidence Intervals

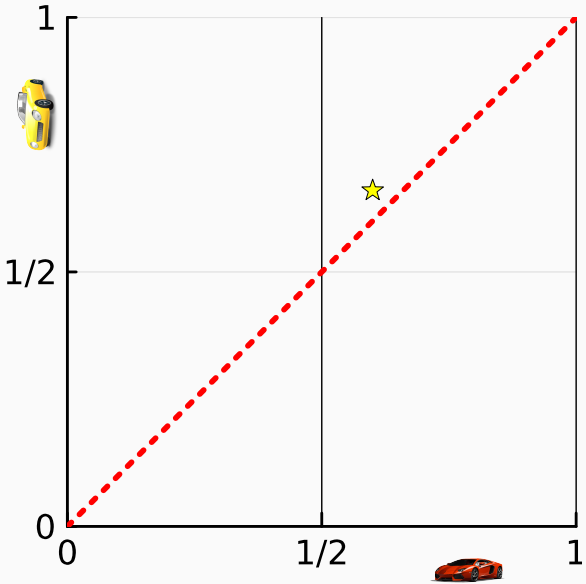
At any point, what do we know?



- Say the success rate is 3/4 and we sample 100 times
- then we will **typically** see a number of successes in the blue centre (67 – 82)
- while a number of successes in either red tail is **weird**
- **Reversing** the logic: if we see ≤ 66 successes or ≥ 83 successes, we **learn** that the success rate is NOT 3/4.

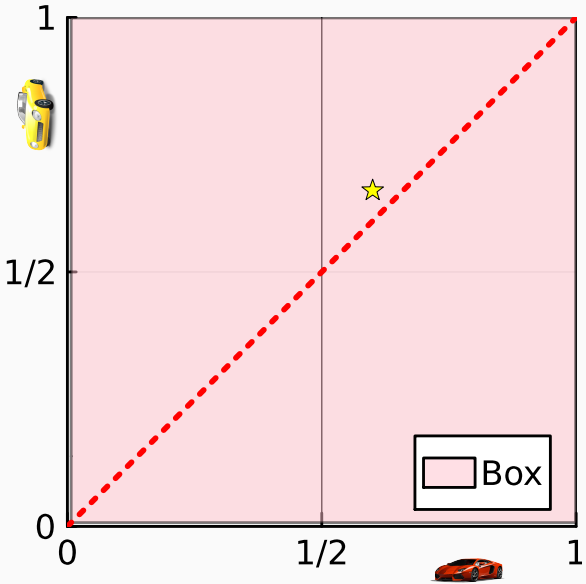
Time 0





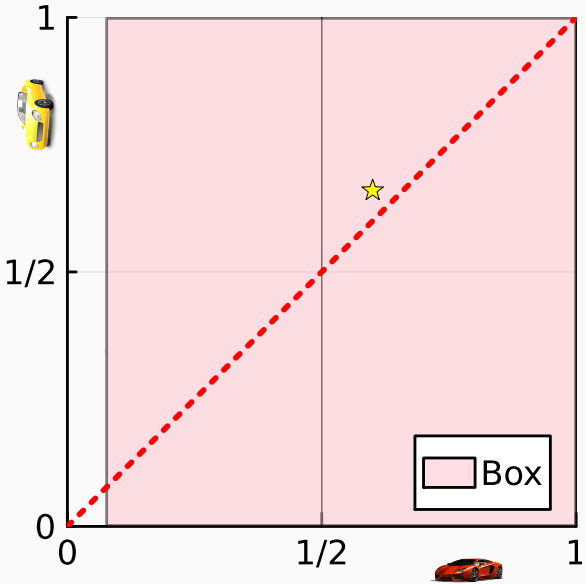
Time 1

	1	0
	1	0



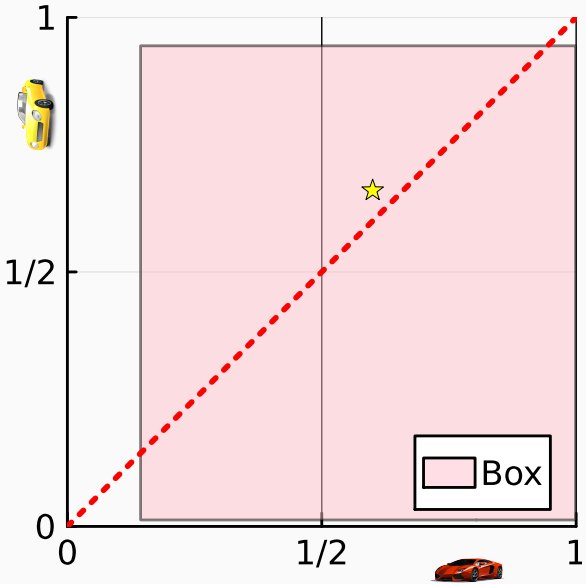
Time 2

	2	0
	1	1



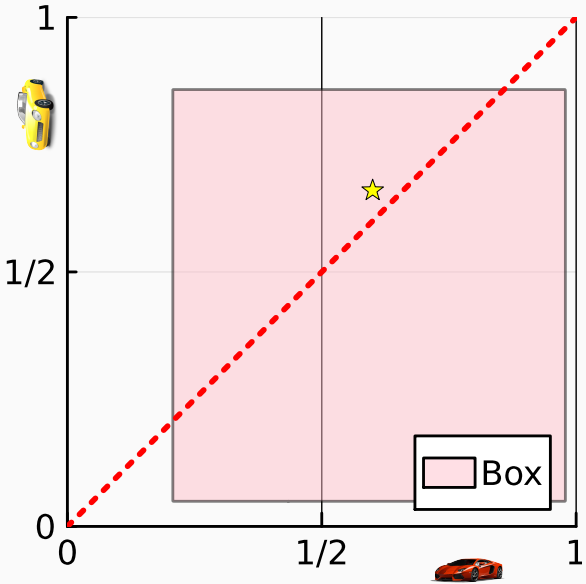
Time 5

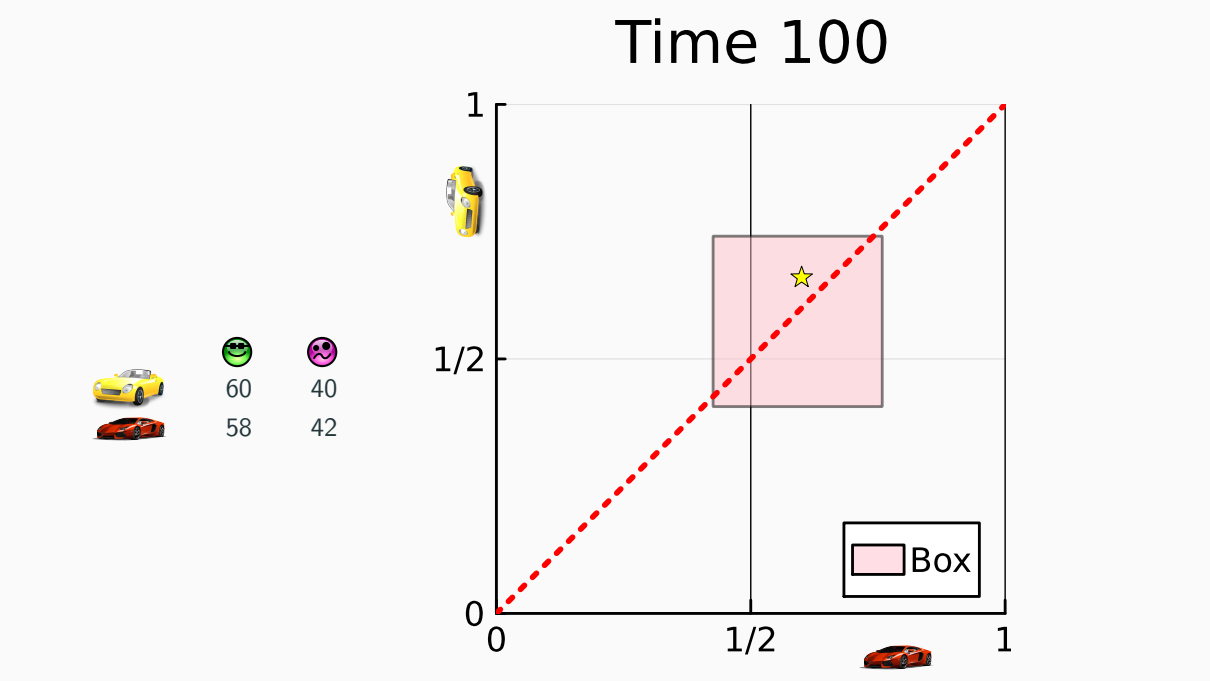
4	1
2	3



Time 10

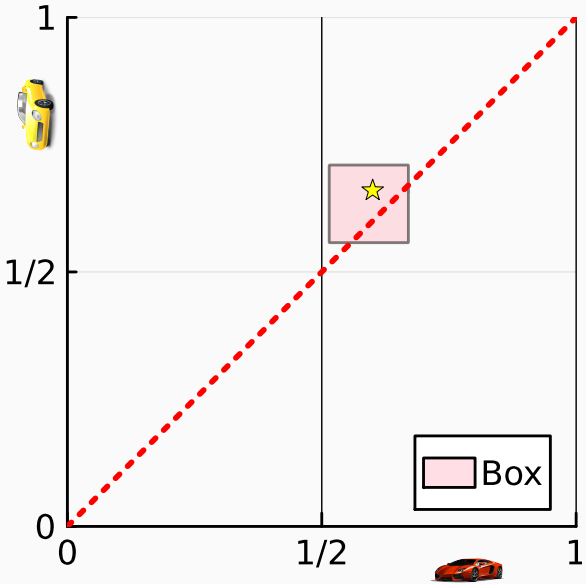
7	3
4	6




Time 100
1
1/2
0
0
1/2
1
Box
60
40
58
42

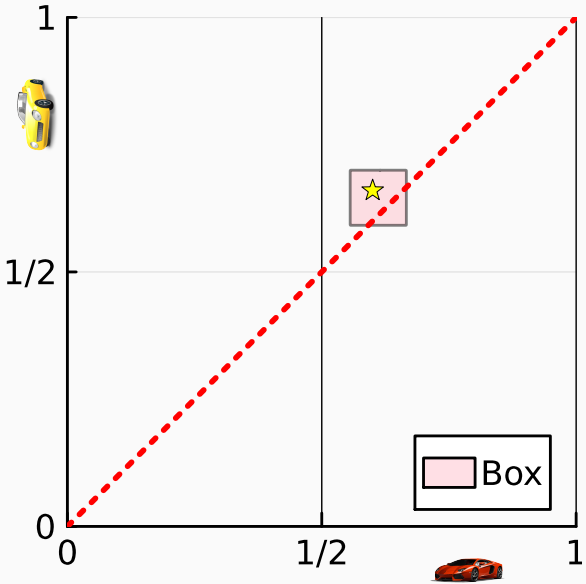
Time 500

	297	203
	318	182

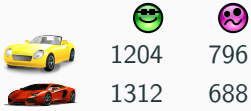


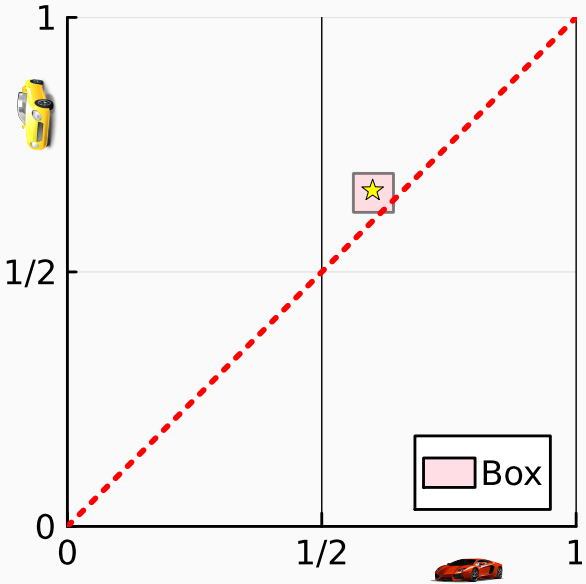
Time 1000

612	388
647	353

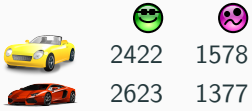


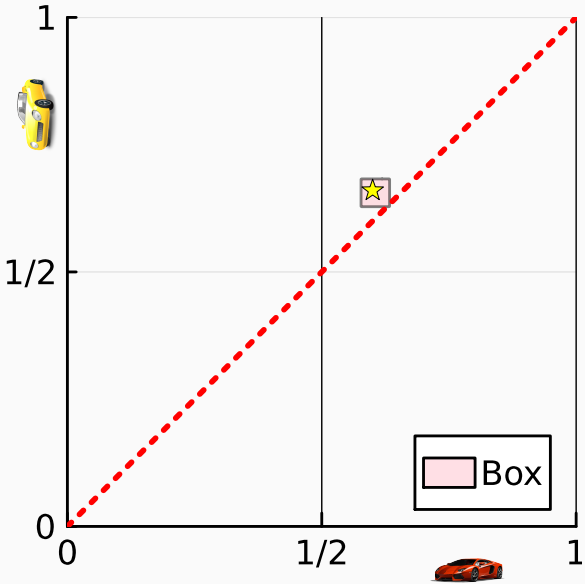
Time 2000



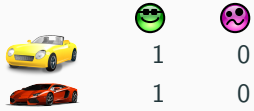


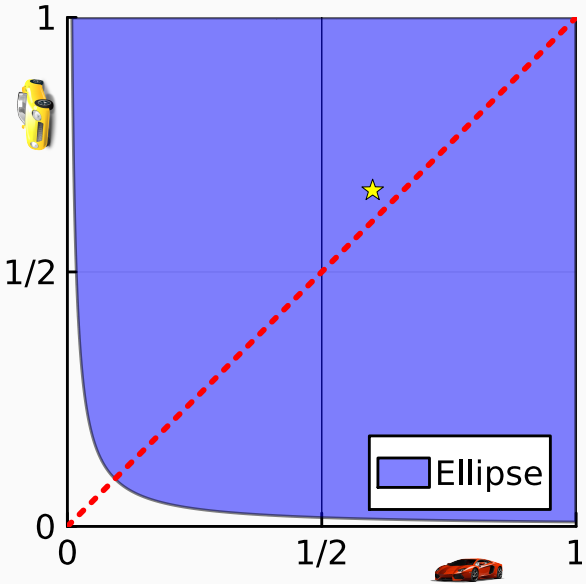
Time 4000





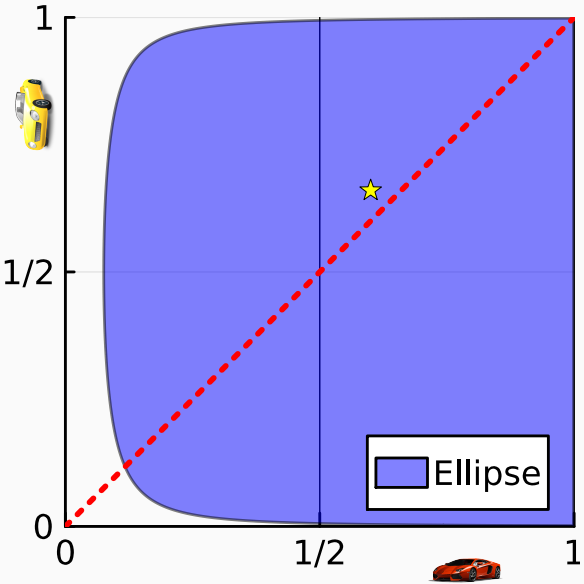
Time 1





Time 2

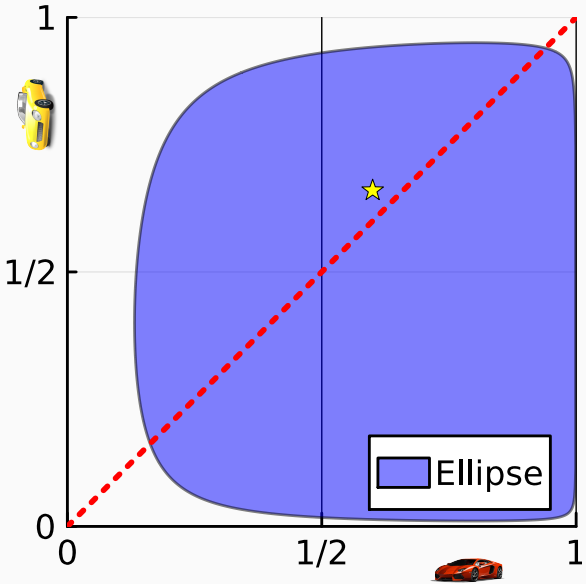
	2	0
	1	1



Time 5

4 1

2 3

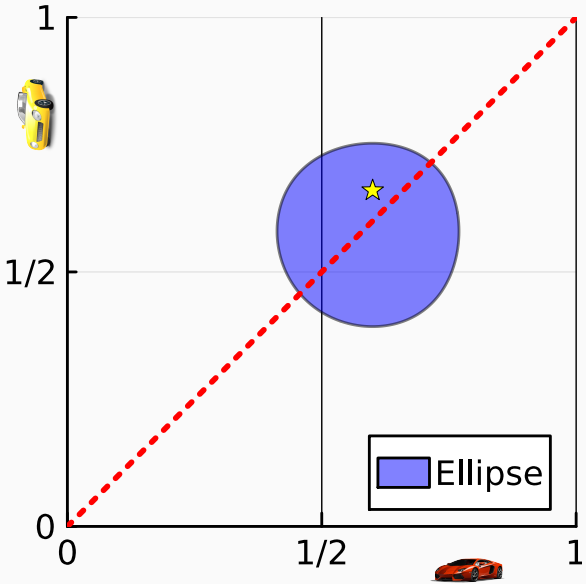


Time 10

	7	3
	4	6

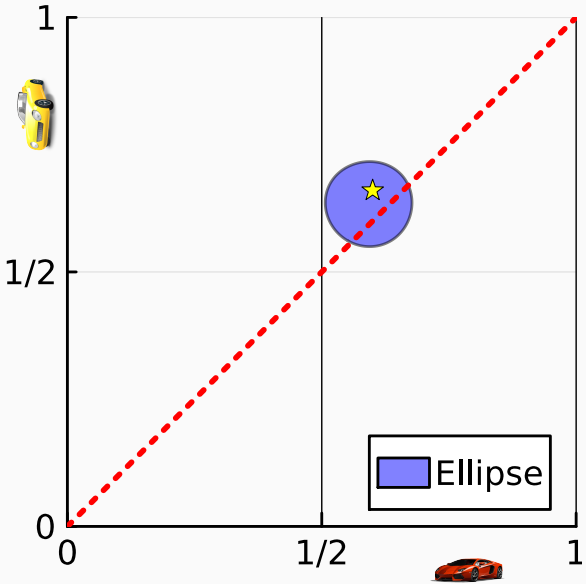
Time 100


	60	40
	58	42



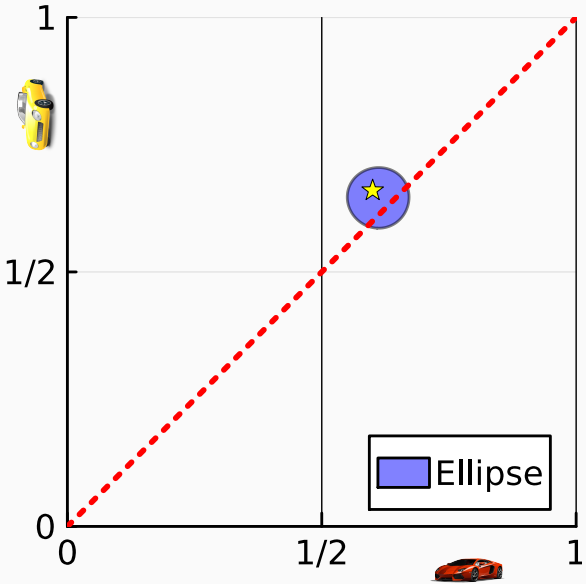
Time 500

	297	203
	318	182

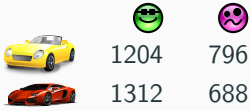


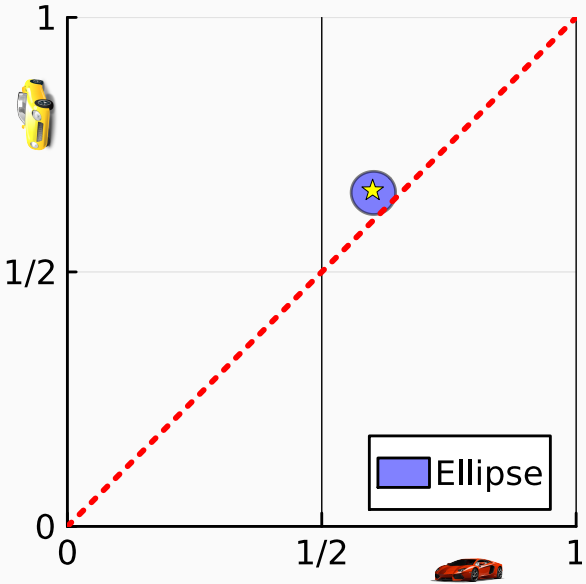
Time 1000


612	388
647	353

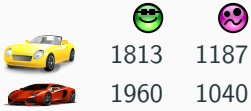


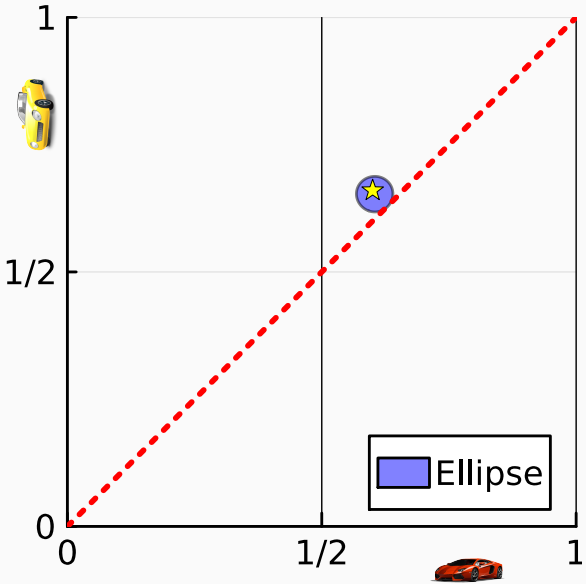
Time 2000



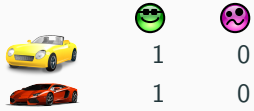


Time 3000

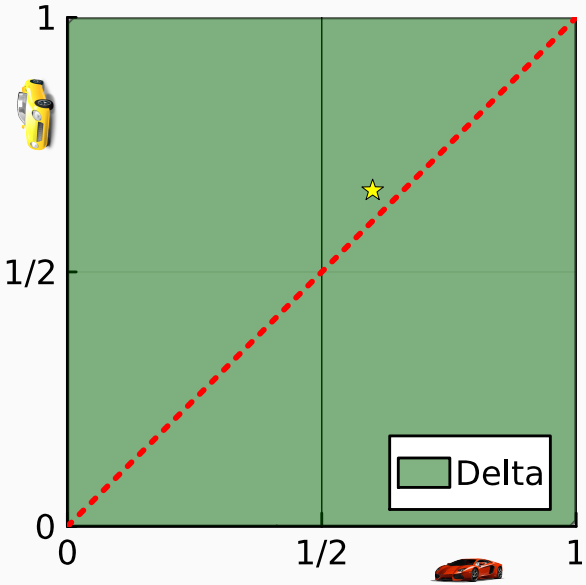




Time 1

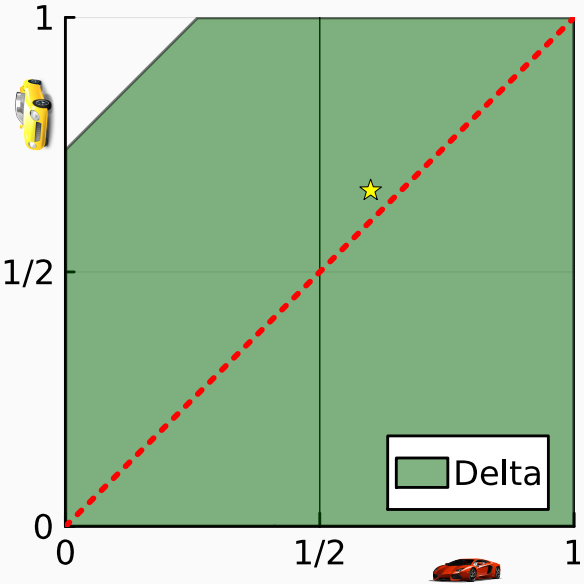


	1	0
	1	0



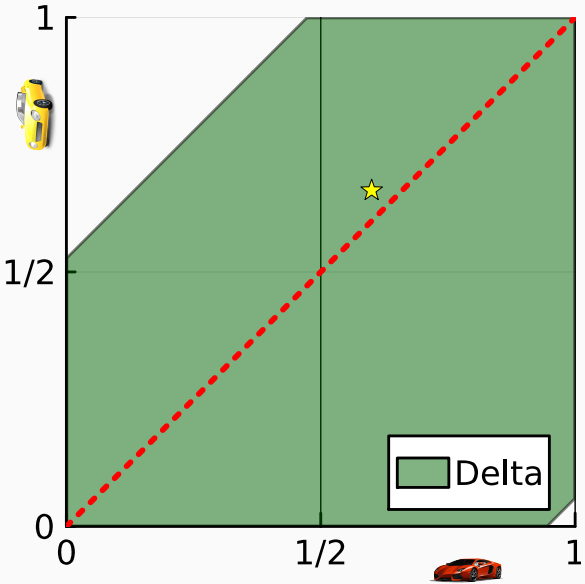
Time 2

	2	0
	1	1



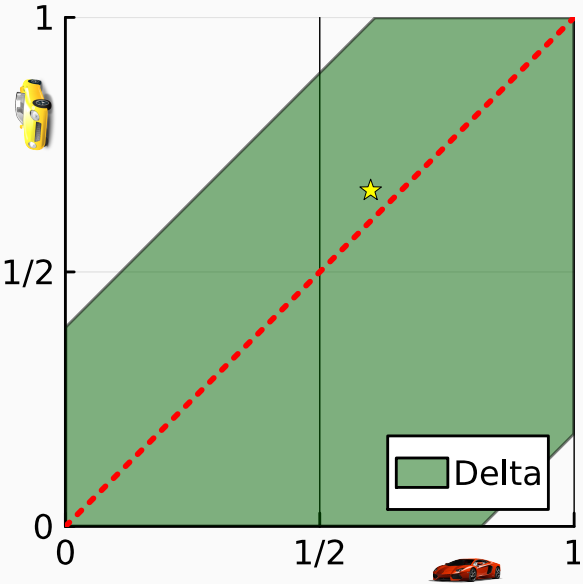
Time 5

	4	1
	2	3



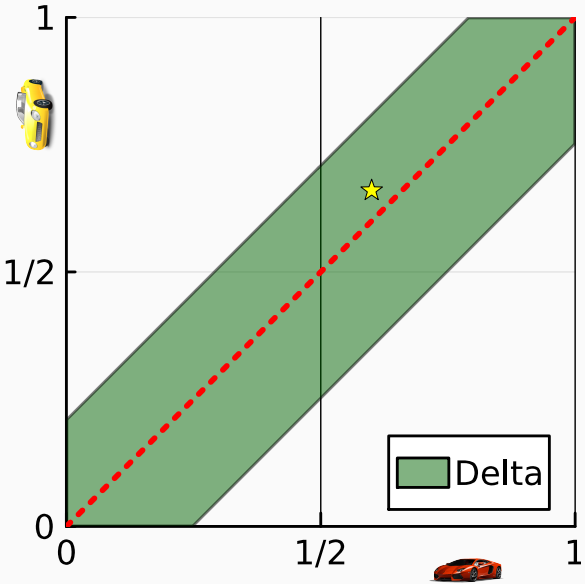
Time 10

	7	3
	4	6



Time 100

60	40
58	42

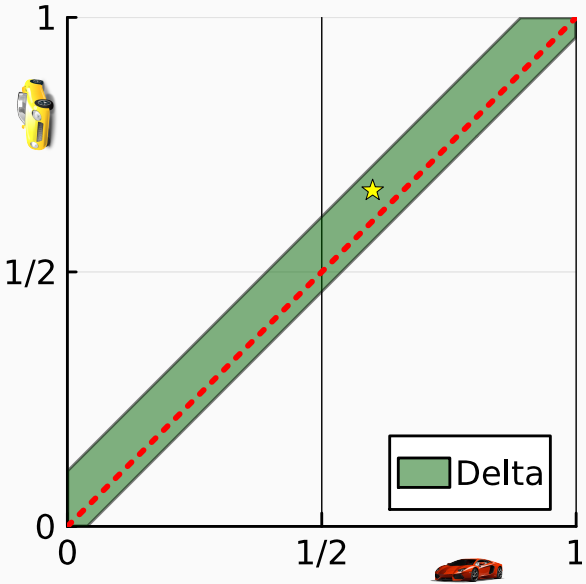


Time 500

	297	203
	318	182

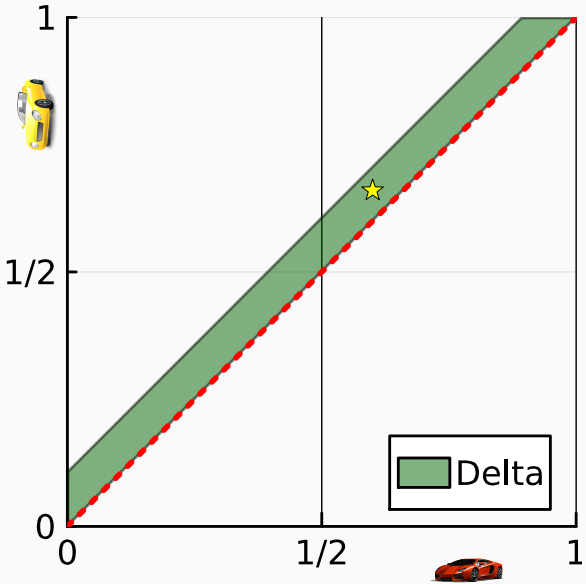
Time 1000

612	388
647	353

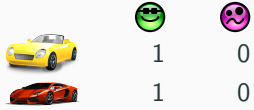


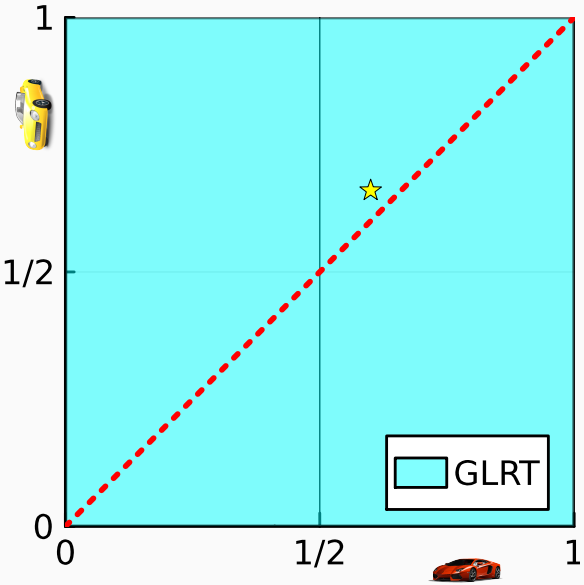
Time 2000

	1204	796
	1312	688



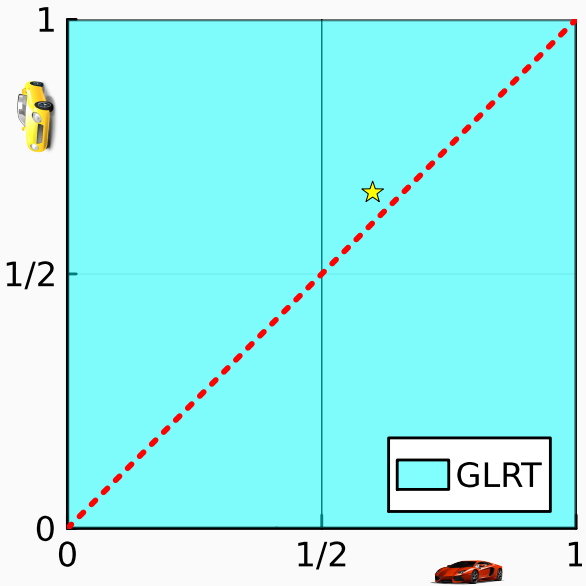
Time 1





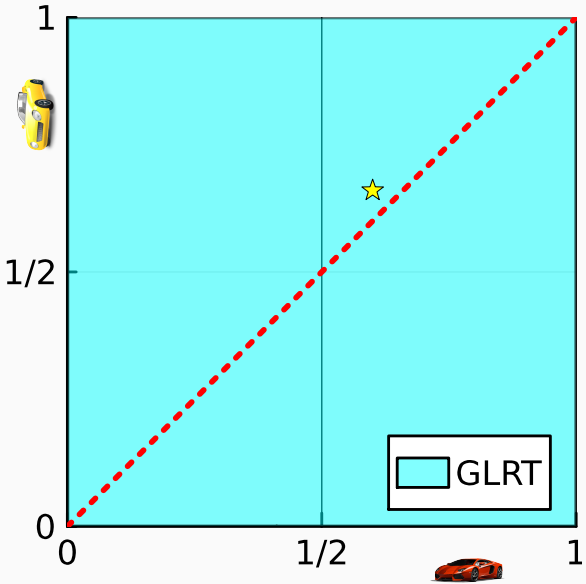
Time 2

	2	0
	1	1



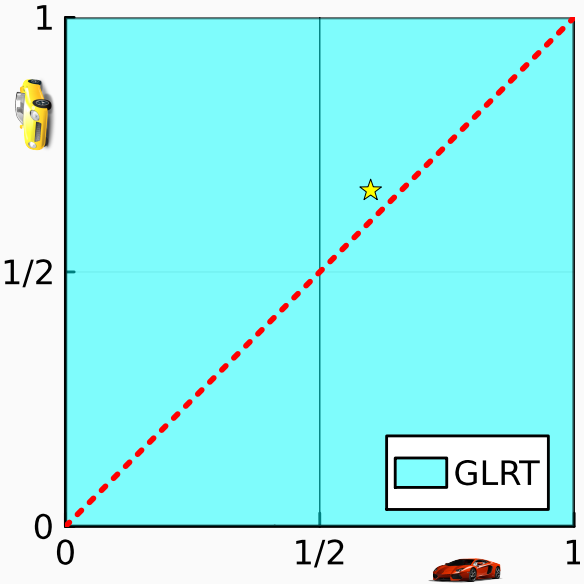
Time 5

4	1
2	3



Time 10

	7	3
	4	6



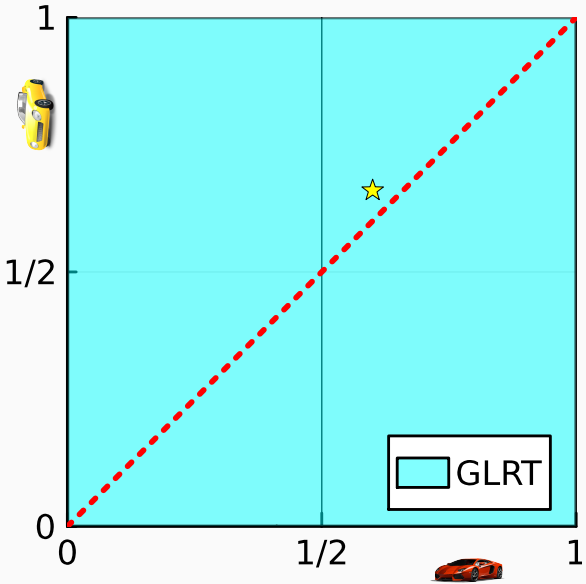
Time 100

60

40

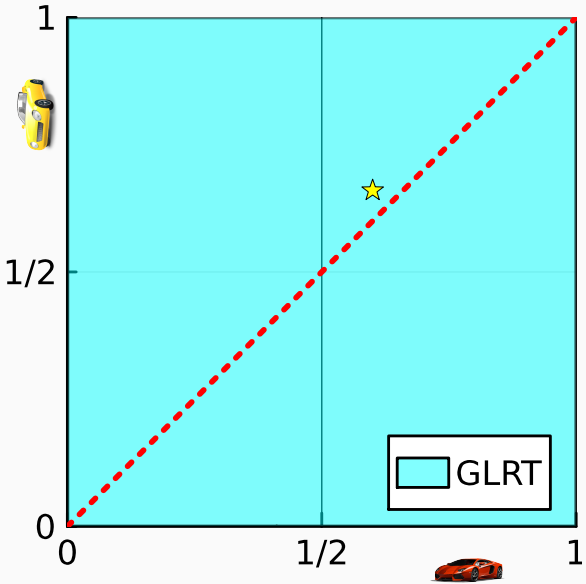
58

42



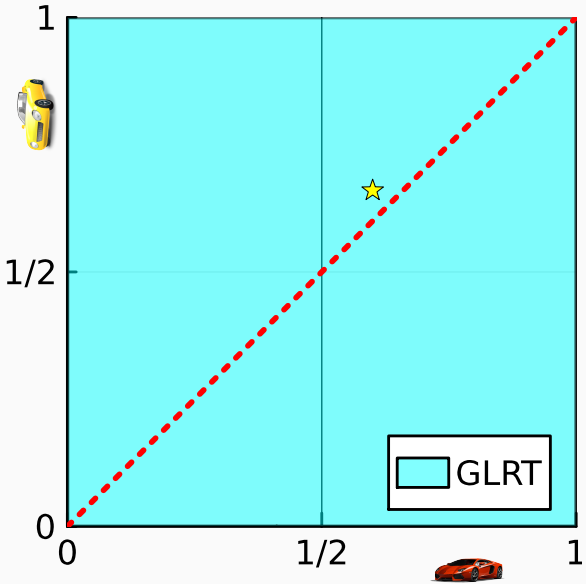
Time 500

	297	203
	318	182

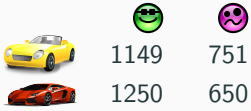


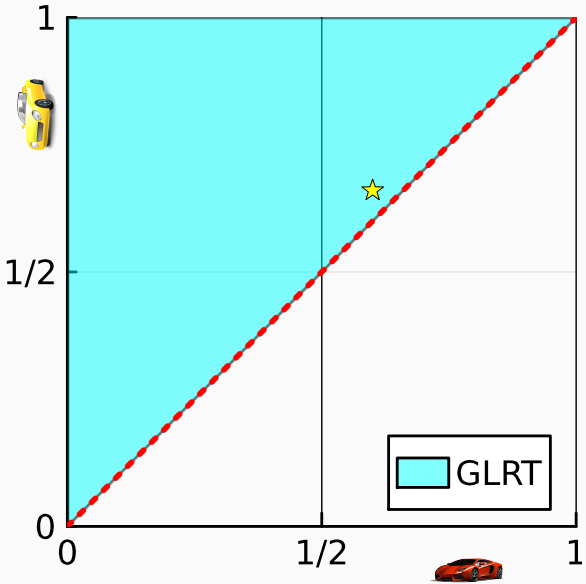
Time 1000

	612	388
	647	353

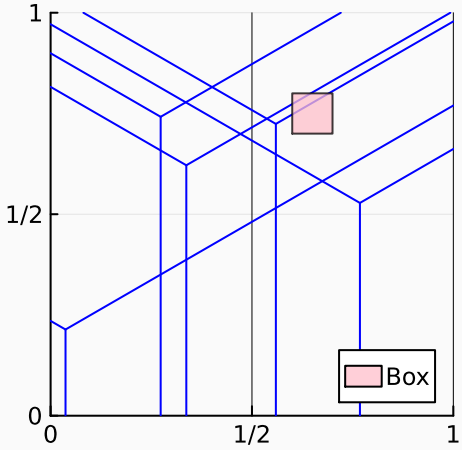


Time 1900





Bespoke Confidence Sequences



How to create problem-aware confidence sequences?

- Promising techniques
  - GLRT
  - martingales / e-variables
- **Challenges**
  - Deviation inequalities (non-asymptotic)
  - Computation

Q2: What should we do

Increased Realism: Different Environments

Versions




Environments





Outcomes


Updated World Model

Known natural environment mix

$\mathbb{P}$  $= 40\%$

$\mathbb{P}$  $= 30\%$

$\mathbb{P}$  $= 15\%$

$\mathbb{P}$  $= 10\%$

$\mathbb{P}$  $= 5\%$

Unknown crash rates (☹/10.000)

					
	10	20	30	40	50
	30	20	20	40	30
	3	3	3	3	400
	40	18	14	10	13

Updated World Model

Known natural environment mix

$\mathbb{P}$() = 40%

$\mathbb{P}$() = 30%

$\mathbb{P}$() = 15%

$\mathbb{P}$() = 10%

$\mathbb{P}$() = 5%

Unknown crash rates (☹/10.000)

					
	10	20	30	40	50
	30	20	20	40	30
	3	3	3	3	400
	40	18	14	10	13

Together these determine the best version on average,  .

Learning the best Version by Exploring

Learning the best Version by Exploring

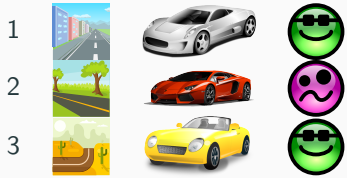
1   

Learning the best Version by Exploring

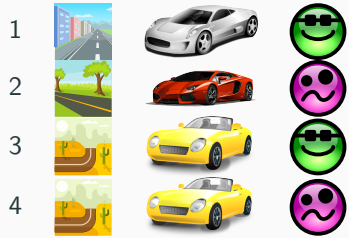
1

2

Learning the best Version by Exploring

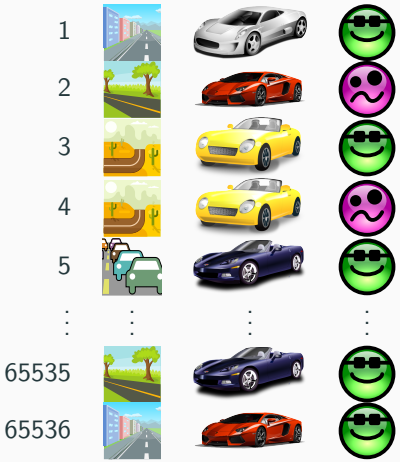


Learning the best Version by Exploring

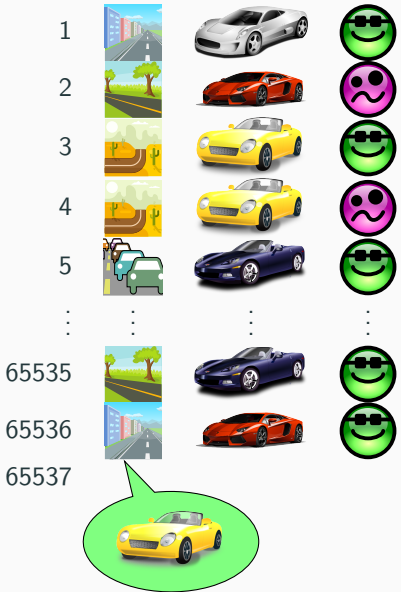


Learning the best Version by Exploring

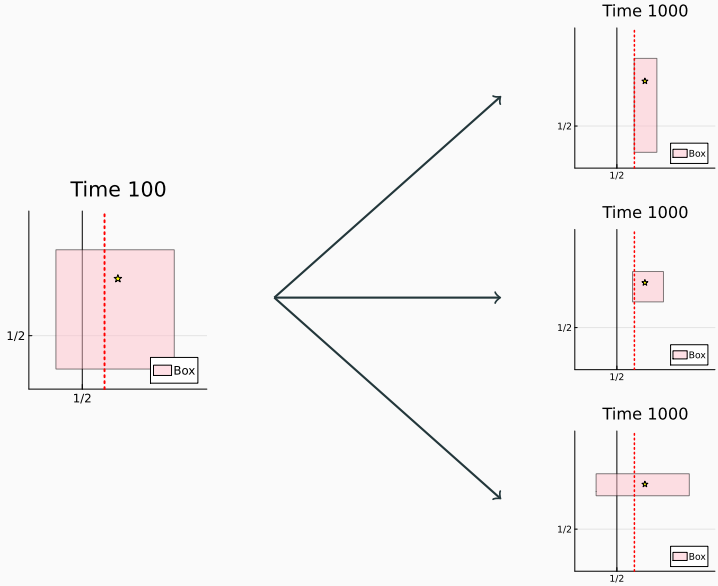
Learning the best Version by Exploring



Learning the best Version by Exploring



Why does Adaptive Exploration Help?



What can Optimal Exploration Achieve?

At $\delta = 5\%$ confidence, need about 1 million samples:

$\times 1000$					
	161	171	104	82	44
	21	12	7	7	2
	77	59	30	20	117
	46	21	9	5	3

These numbers depend on the unknown crash rates. Only way to match these is to **learn about them** on the fly.

Efficiency Gain

- In my example, **adaptive** exploration reduces experiments by factor 3 (!) compared to **batch**.
  - Focus on discriminating better versions
  - Not smart to drive with the natural mix of environments.

The Future

- Main challenges
  - Optimal behaviour determined by unknown, **non-convex** saddle point
  - Statistics, **Optimisation**
- The dream: automate strategy design

Personal Journey

My deep gratitude

- Peter Grünwald and Paul Vitányi (UvA / ILLC / CWI)
- Manfred Warmuth (UC Santa Cruz)
- Volodya Vovk (Royal Holloway)
- Peter Bartlett (Berkeley and Brisbane)
- Johannes Schmidt-Hieber (University of Twente)

Online Learning Bandits Pure Exploration Optimisation Zero-sum games Testing Martingales
Probability Algorithms

Ik heb gezegd!