

# Space-Efficient Scheduling of Stochastically Generated Tasks

Tomáš Brázdil<sup>1</sup>   Javier Esparza<sup>2</sup>  
Stefan Kiefer<sup>3</sup>   Michael Luttenberger<sup>2</sup>

<sup>1</sup>Masaryk University, Brno (Czech Republic)

<sup>2</sup>TU München (Germany)

<sup>3</sup>University of Oxford (UK)

ICALP, Bordeaux (July 2010)

We study **tasks** that stochastically **generate new tasks**.  
The execution of a task can generate subtasks.

## Examples:

- **Divide-and-conquer Algorithms:**  
Either solve the problem directly or solve sub-instances.
- **Multi-Threaded Programs:**  
Either terminate, or spawn new thread, or none of that.

We study **tasks** that stochastically **generate new tasks**.  
The execution of a task can generate subtasks.

## Examples:

- **Divide-and-conquer Algorithms:**  
Either solve the problem directly or solve sub-instances.
- **Multi-Threaded Programs:**  
Either terminate, or spawn new thread, or none of that.

How much memory is needed?

# Task Systems

Formally, task systems are like context-free grammars.

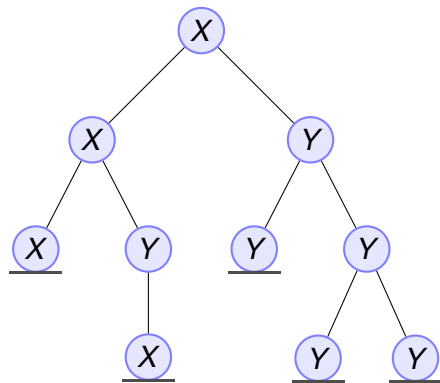
$$X \hookrightarrow XY$$

$$X \hookrightarrow \varepsilon$$

$$Y \hookrightarrow YY$$

$$Y \hookrightarrow X$$

$$Y \hookrightarrow \varepsilon$$



# Task Systems

Formally, task systems are like context-free grammars.

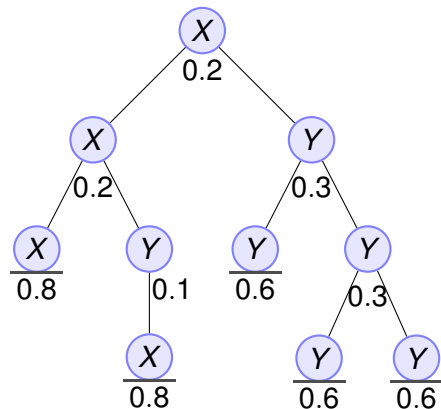
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$



# Task Systems

Formally, task systems are like context-free grammars.

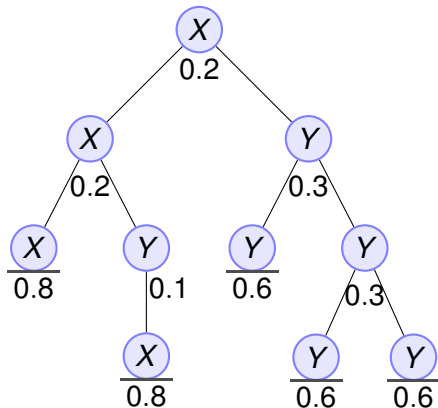
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$



Probability of this tree:

$$0.2 \cdot 0.2 \cdot 0.3 \cdot 0.8 \cdot 0.1 \cdot 0.6 \cdot 0.3 \cdot 0.8 \cdot 0.6 \cdot 0.6$$

# A Derivation

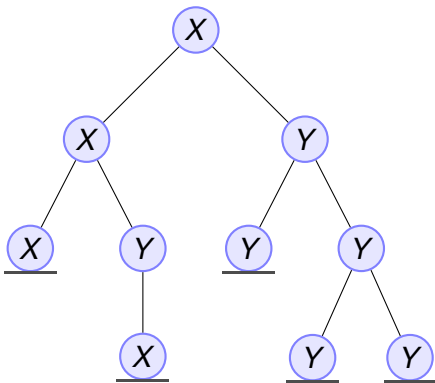
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



# A Derivation

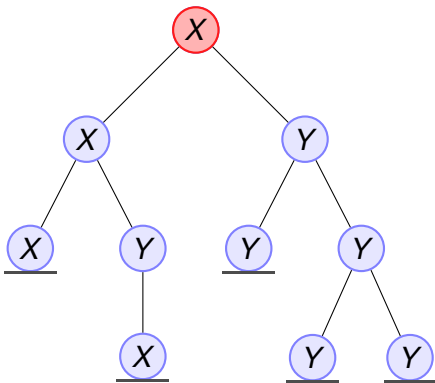
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



X



# A Derivation

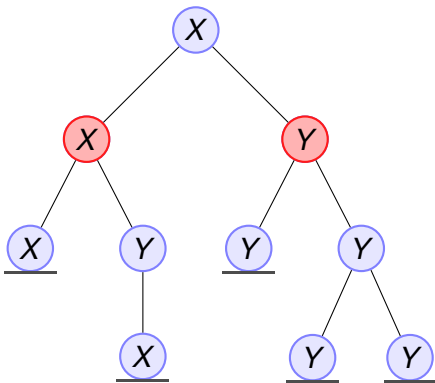
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY$

# A Derivation

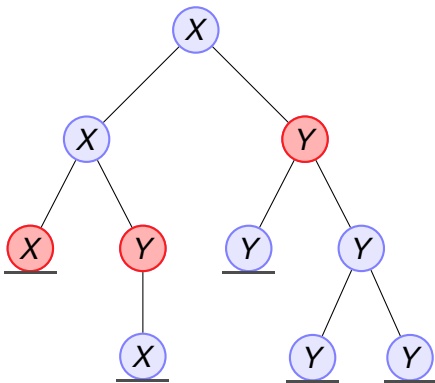
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow \mathbf{XXY}$

# A Derivation

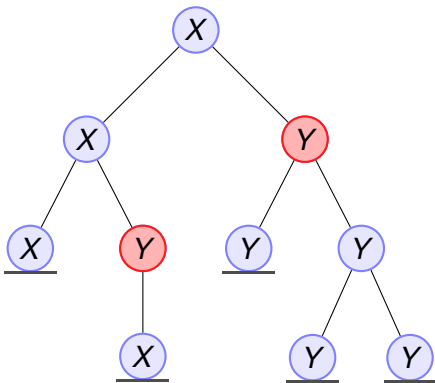
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XXY \Rightarrow YY$

# A Derivation

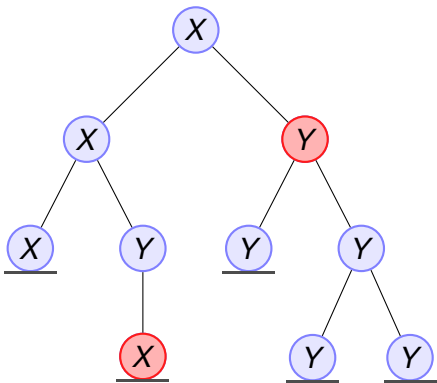
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XXY \Rightarrow YY \Rightarrow XY$

# A Derivation

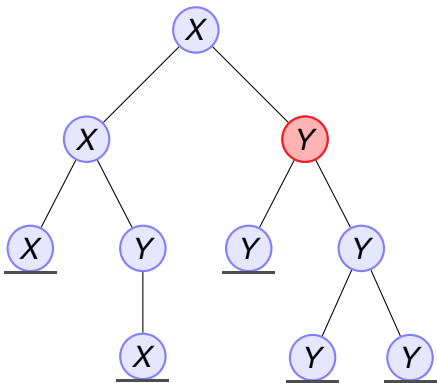
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XXY \Rightarrow YY \Rightarrow XY \Rightarrow Y$

# A Derivation

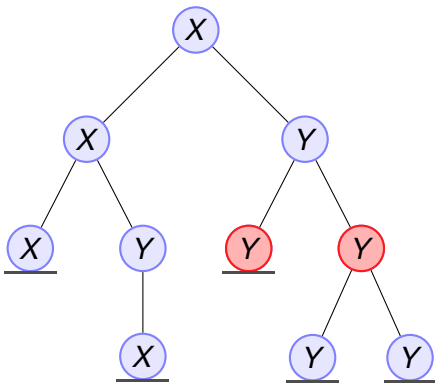
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XXY \Rightarrow YY \Rightarrow XY \Rightarrow Y \Rightarrow YY$

# A Derivation

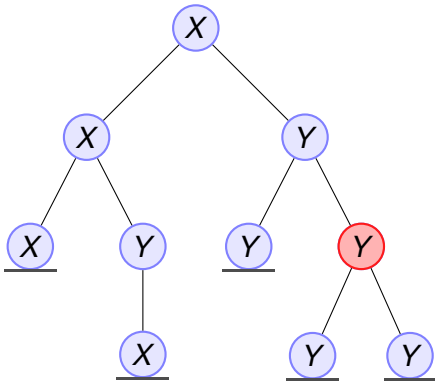
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XXY \Rightarrow YY \Rightarrow XY \Rightarrow Y \Rightarrow YY \Rightarrow Y$

# A Derivation

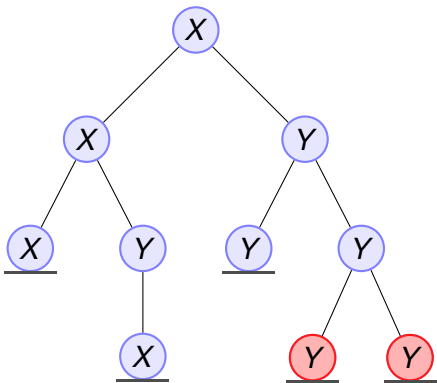
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XXY \Rightarrow YY \Rightarrow XY \Rightarrow Y \Rightarrow YY \Rightarrow Y \Rightarrow YY$



# A Derivation

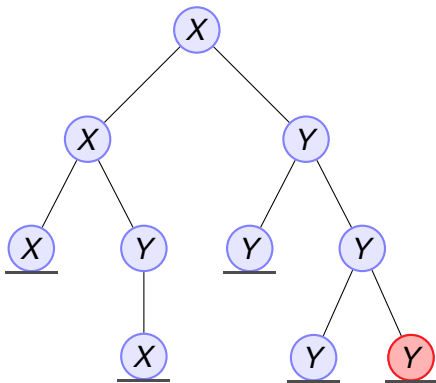
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XXY \Rightarrow YY \Rightarrow XY \Rightarrow Y \Rightarrow YY \Rightarrow Y \Rightarrow YY \Rightarrow Y$

# A Derivation

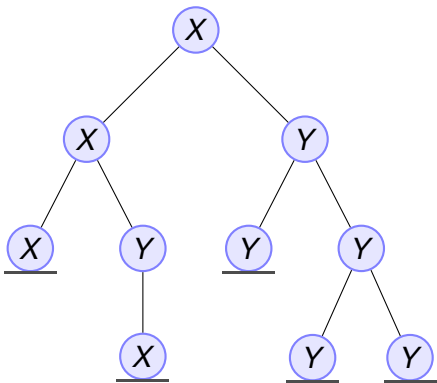
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XXY \Rightarrow YY \Rightarrow XY \Rightarrow Y \Rightarrow YY \Rightarrow Y \Rightarrow YY \Rightarrow Y \Rightarrow \varepsilon$

# A Derivation

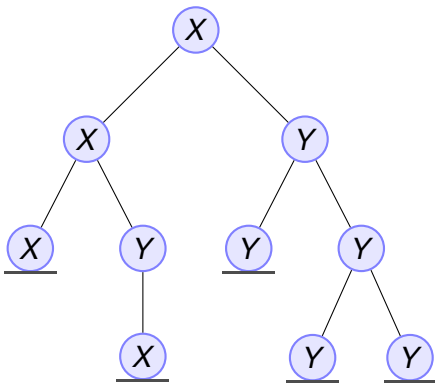
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XXY \Rightarrow YY \Rightarrow XY \Rightarrow Y \Rightarrow YY \Rightarrow Y \Rightarrow YY \Rightarrow Y \Rightarrow \varepsilon$

Time: 10 (number of nodes)

Space: 3



# Another Derivation

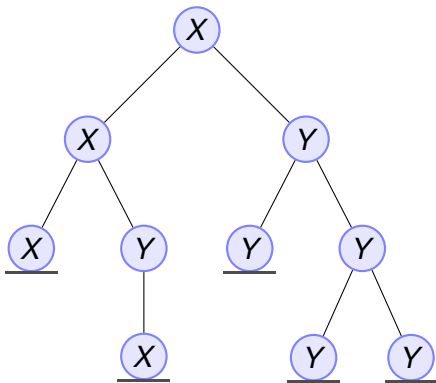
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



# Another Derivation

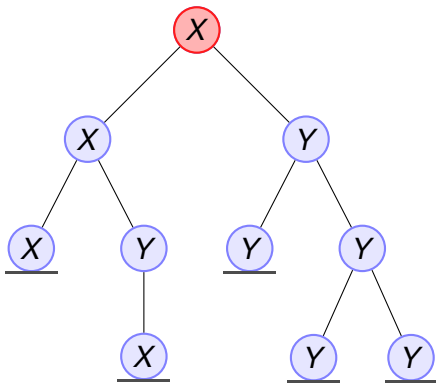
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



X

# Another Derivation

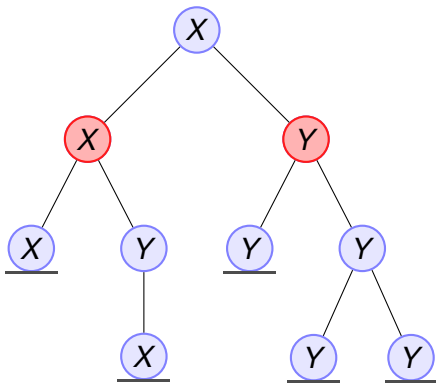
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY$

# Another Derivation

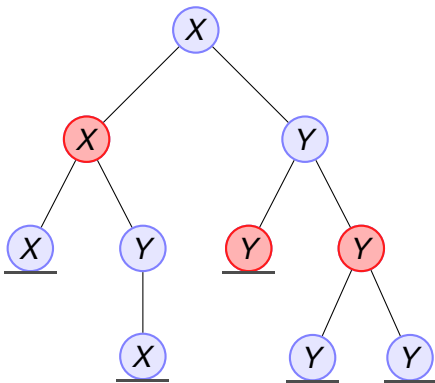
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow \color{red}{XYY}$



# Another Derivation

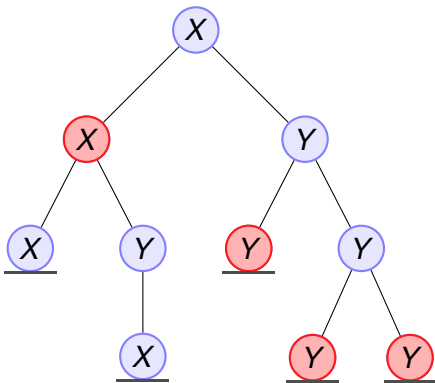
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XYY \Rightarrow \color{red}{XYYY}$

# Another Derivation

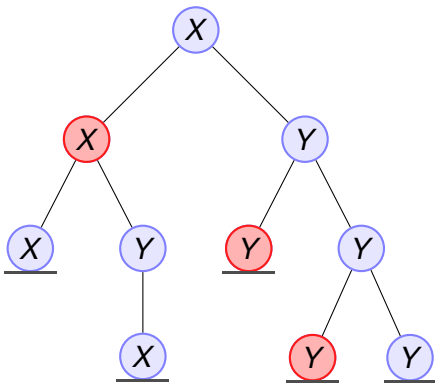
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XYY \Rightarrow XYYY \Rightarrow \color{red}{XYY}$

# Another Derivation

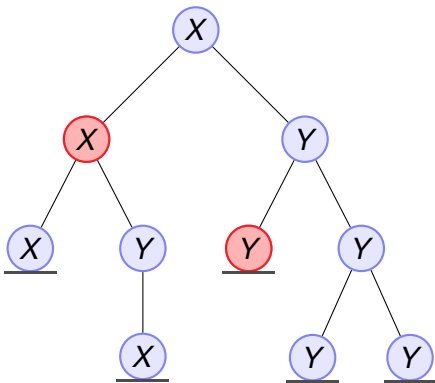
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XYY \Rightarrow XYYY \Rightarrow XYY \Rightarrow XY$

# Another Derivation

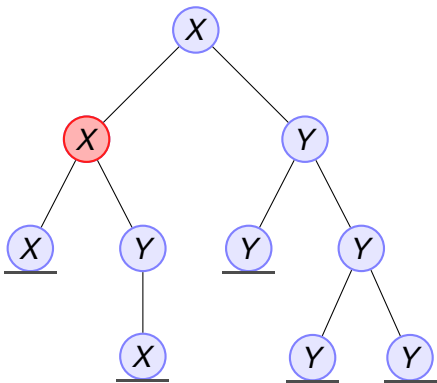
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XYY \Rightarrow XYYY \Rightarrow XYY \Rightarrow XY \Rightarrow X$

# Another Derivation

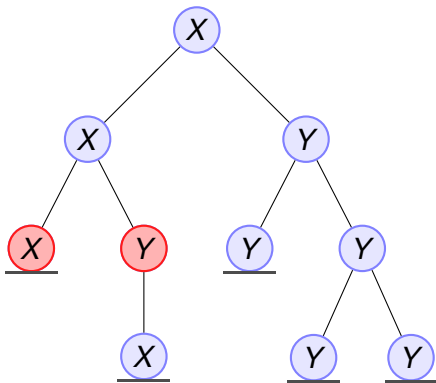
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XYY \Rightarrow XYYY \Rightarrow XYY \Rightarrow XY \Rightarrow X \Rightarrow XY$



# Another Derivation

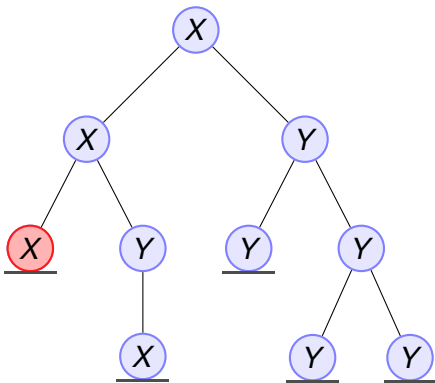
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XYY \Rightarrow XYYY \Rightarrow XYY \Rightarrow XY \Rightarrow X \Rightarrow XY \Rightarrow XX \Rightarrow X$





# Another Derivation

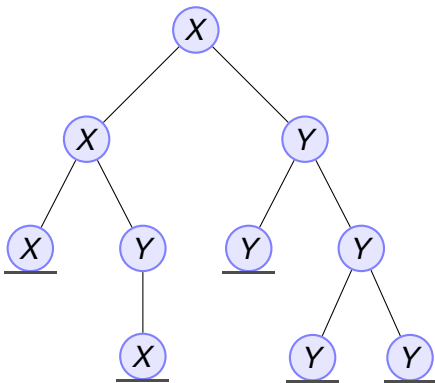
$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$

$Y \hookrightarrow YY$

$Y \hookrightarrow X$

$Y \hookrightarrow \varepsilon$



$X \Rightarrow XY \Rightarrow XYY \Rightarrow XYYY \Rightarrow XYY \Rightarrow XY \Rightarrow X \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$

Time: 10 (number of nodes)

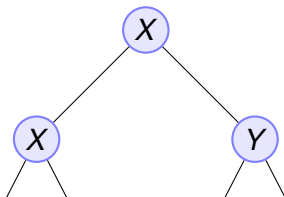
Space: 4



# Another Derivation

$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$



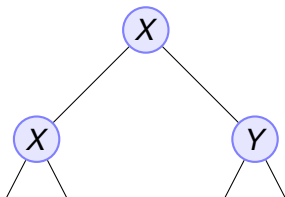
Each finite tree has its

- Probability (does not depend on scheduler)
- Time (does not depend on scheduler)
- Space (**depends on scheduler**)

# Another Derivation

$X \hookrightarrow XY$

$X \hookrightarrow \varepsilon$



Each finite tree has its

- Probability (does not depend on scheduler)
- Time (does not depend on scheduler)
- Space (**depends on scheduler**)

Raises questions like:

- What is the **expected** time and the expected space?
- How are these random variables **distributed**?

Time has been studied before. We focus on space.

# Related Work: Branching Processes

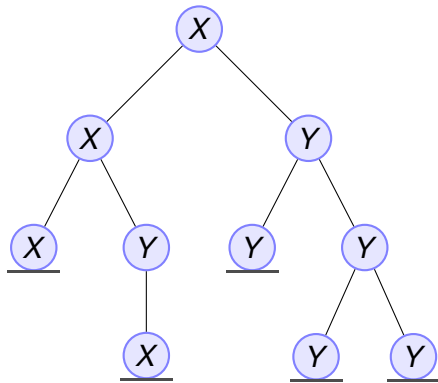
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

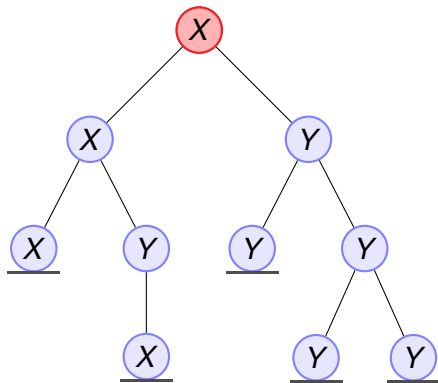
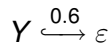
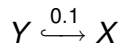
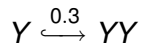
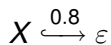
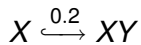
$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

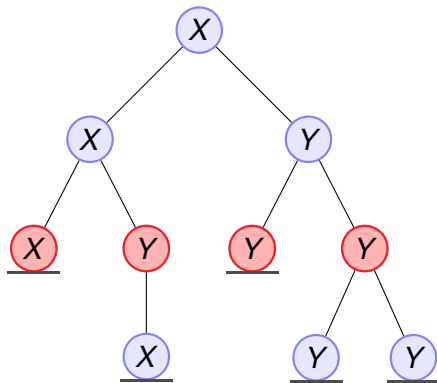
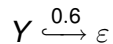
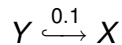
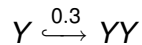
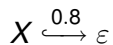
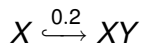


# Related Work: Branching Processes





# Related Work: Branching Processes





# Related Work: Branching Processes

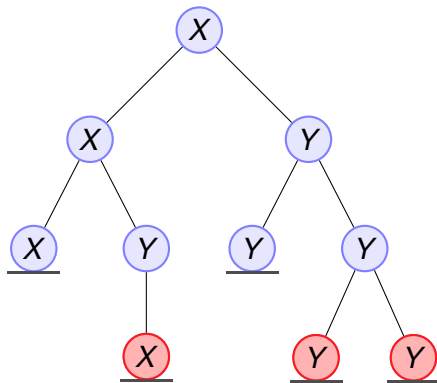
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$



# Related Work: Branching Processes

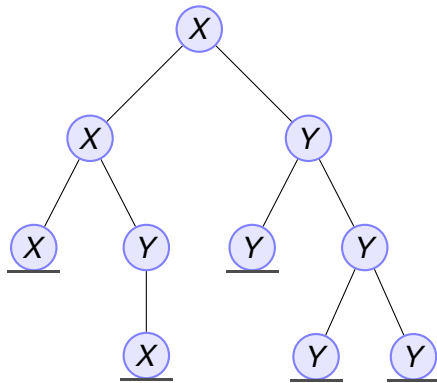
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$





# Termination Probability

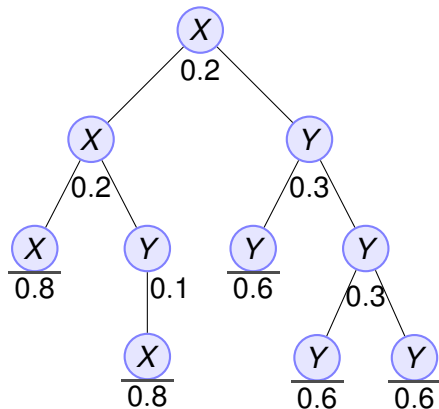
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$



Each tree has its probability.

The sum of these probabilities is the “termination probability”.

Is it always 1?

# Termination Probability and the function $f$

A task system induces a vector  $f(\mathbf{x})$ .

For our example:  $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$  and  $f(\mathbf{x}) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix}$  with

$$\left. \begin{array}{l} X \xrightarrow{0.2} XY \\ X \xrightarrow{0.8} \varepsilon \end{array} \right\} f_x(x, y) = 0.2xy + 0.8$$

$$\left. \begin{array}{l} Y \xrightarrow{0.3} YY \\ Y \xrightarrow{0.1} X \\ Y \xrightarrow{0.6} \varepsilon \end{array} \right\} f_y(x, y) = 0.3y^2 + 0.1x + 0.6$$

# Termination Probability and the function $f$

A task system induces a vector  $f(\mathbf{x})$ .

For our example:  $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$  and  $f(\mathbf{x}) = \begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix}$  with

$$\left. \begin{array}{l} X \xrightarrow{0.2} XY \\ X \xrightarrow{0.8} \varepsilon \end{array} \right\} f_x(x, y) = 0.2xy + 0.8$$

$$\left. \begin{array}{l} Y \xrightarrow{0.3} YY \\ Y \xrightarrow{0.1} X \\ Y \xrightarrow{0.6} \varepsilon \end{array} \right\} f_y(x, y) = 0.3y^2 + 0.1x + 0.6$$

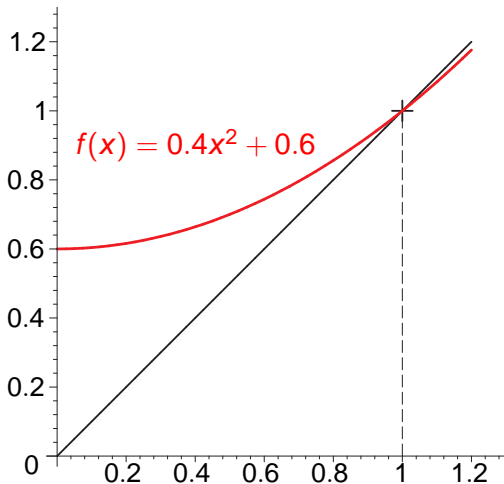
Proposition (well-known, see [Harris])

*The termination probability is the (first component of the) **least fixed point of  $f$** .*

# The function $f$

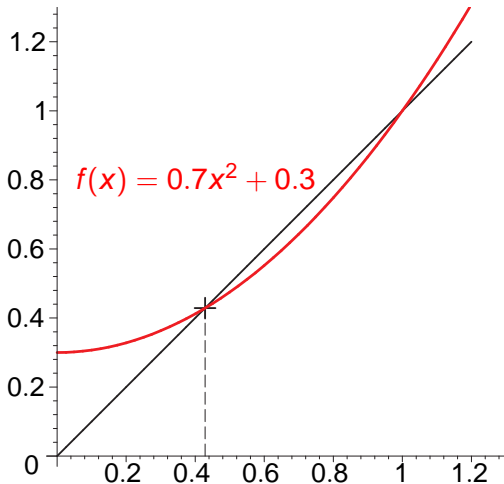
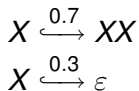
The **subcritical** case:  $\begin{cases} \text{termination probability} & = 1 \\ \text{expected time} & = \text{finite} \end{cases}$

$$\begin{aligned} X &\xrightarrow{0.4} XX \\ X &\xrightarrow{0.6} \varepsilon \end{aligned}$$



# The function $f$

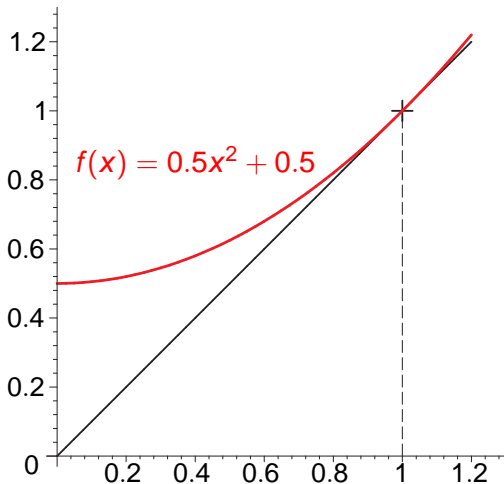
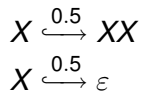
The **supercritical** case:  $\begin{cases} \text{termination probability} < 1 \\ \text{expected time} = \infty \end{cases}$





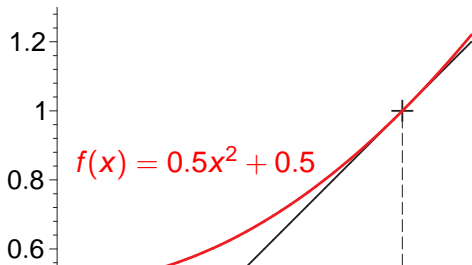
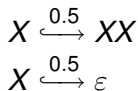
# The function $f$

The **critical** case:  $\left\{ \begin{array}{l} \text{termination probability} = 1 \\ \text{expected time} = \infty \end{array} \right.$



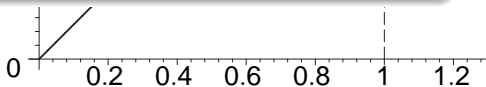
# The function $f$

The **critical** case:  $\begin{cases} \text{termination probability} & = 1 \\ \text{expected time} & = \infty \end{cases}$



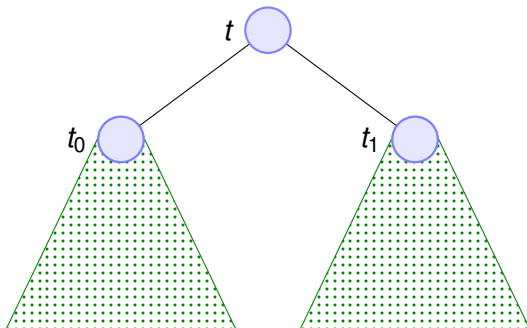
## Assumption

We assume termination probability = 1 in the following, i.e., subcritical or critical.



# Optimal Scheduling

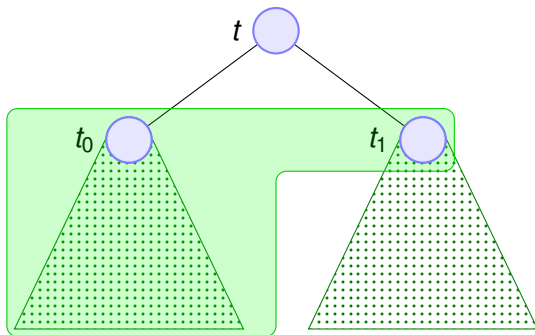
Given a tree  $t$  with two children  $t_0, t_1$ .  
What is the optimal scheduling?



$$S^{op}(t) =$$

# Optimal Scheduling

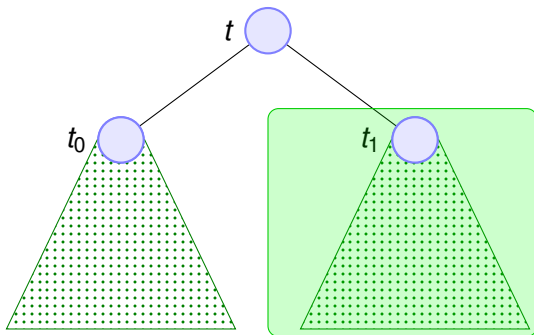
Given a tree  $t$  with two children  $t_0, t_1$ .  
What is the optimal scheduling?



$$S^{op}(t) = \left\{ \max \left\{ S^{op}(t_0) + 1, S^{op}(t_1) \right\} \right\}$$

# Optimal Scheduling

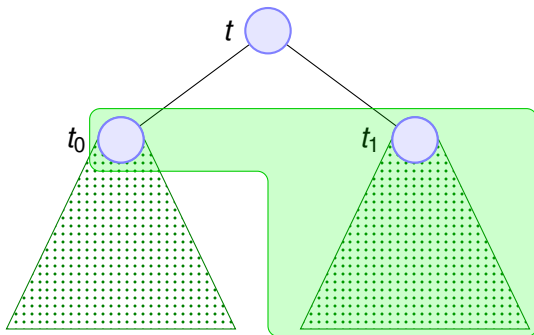
Given a tree  $t$  with two children  $t_0, t_1$ .  
What is the optimal scheduling?



$$S^{op}(t) = \left\{ \max \left\{ S^{op}(t_0) + 1, S^{op}(t_1) \right\} \right\}$$

# Optimal Scheduling

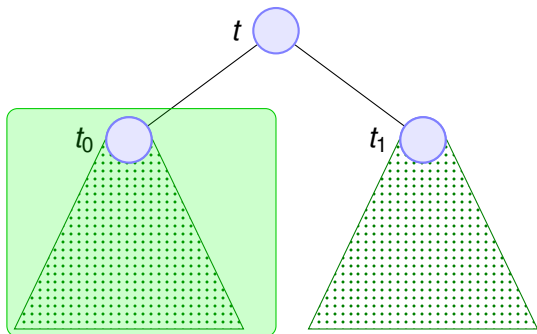
Given a tree  $t$  with two children  $t_0, t_1$ .  
What is the optimal scheduling?



$$S^{op}(t) = \left\{ \begin{array}{l} \max \{ S^{op}(t_0) + 1, S^{op}(t_1) \}, \\ \max \{ S^{op}(t_1) + 1, S^{op}(t_0) \} \end{array} \right\}$$

# Optimal Scheduling

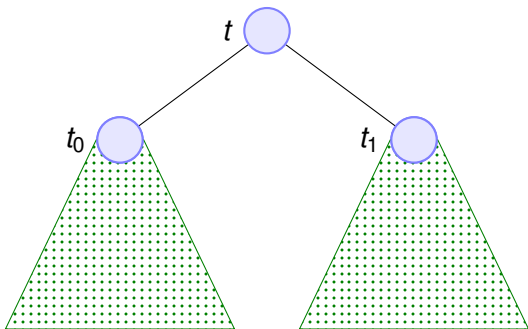
Given a tree  $t$  with two children  $t_0, t_1$ .  
What is the optimal scheduling?



$$S^{op}(t) = \left\{ \begin{array}{l} \max \{ S^{op}(t_0) + 1, S^{op}(t_1) \}, \\ \max \{ S^{op}(t_1) + 1, S^{op}(t_0) \} \end{array} \right\}$$

# Optimal Scheduling

Given a tree  $t$  with two children  $t_0, t_1$ .  
What is the optimal scheduling?

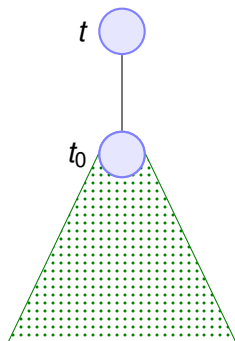


$$S^{op}(t) = \min \left\{ \begin{array}{l} \max \{ S^{op}(t_0) + 1, S^{op}(t_1) \}, \\ \max \{ S^{op}(t_1) + 1, S^{op}(t_0) \} \end{array} \right\}$$



# Optimal Scheduling

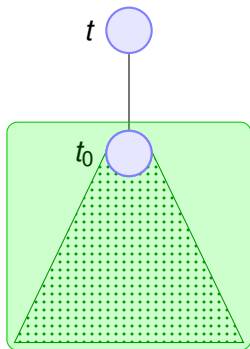
Given a tree  $t$  with just one child  $t_0$ .  
What is the optimal scheduling?



$$S^{op}(t) =$$

# Optimal Scheduling

Given a tree  $t$  with just one child  $t_0$ .  
What is the optimal scheduling?



$$S^{op}(t) = S^{op}(t_0)$$

So we can determine  $S^{op}$  for any tree  $t$ :

$$S^{op}(t) = \begin{cases} \min \left\{ \begin{array}{l} \max\{S^{op}(t_0) + 1, S^{op}(t_1)\}, \\ \max\{S^{op}(t_1) + 1, S^{op}(t_0)\} \end{array} \right\} & \text{if } t \text{ has two children } t_0, t_1 \\ S^{op}(t_0) & \text{if } t \text{ has one child } t_0 \\ 1 & \text{if } t \text{ has no children.} \end{cases}$$

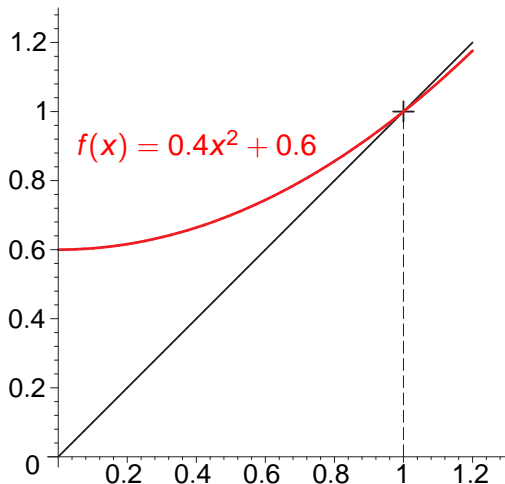
So we can determine  $S^{op}$  for any tree  $t$ :

$$S^{op}(t) = \begin{cases} \min \left\{ \begin{array}{l} \max\{S^{op}(t_0) + 1, S^{op}(t_1)\}, \\ \max\{S^{op}(t_1) + 1, S^{op}(t_0)\} \end{array} \right\} & \text{if } t \text{ has two children } t_0, t_1 \\ S^{op}(t_0) & \text{if } t \text{ has one child } t_0 \\ 1 & \text{if } t \text{ has no children.} \end{cases}$$

What is the distribution of  $S^{op}$ , if trees are randomly generated?

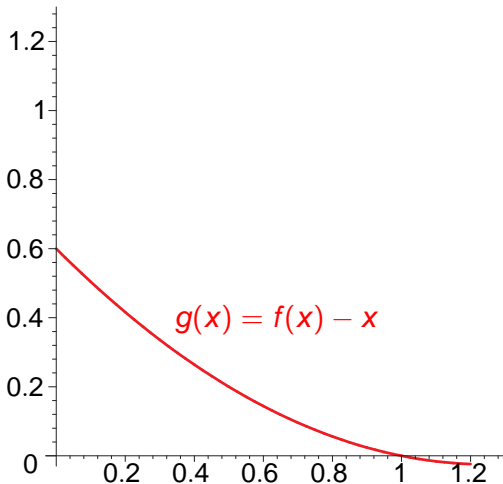
# Newton's Method

Let  $g(x) := f(x) - x$  and apply Newton's method to  $g(x) = 0$ :



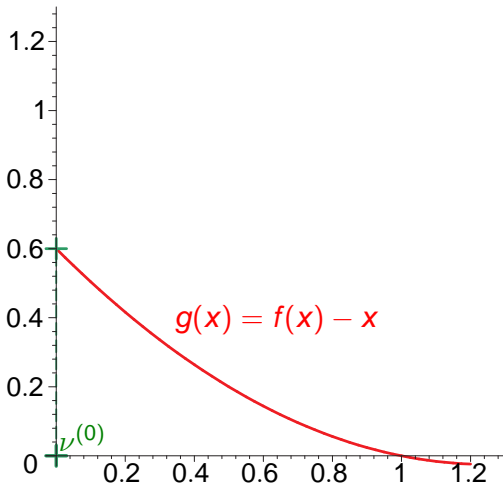
# Newton's Method

Let  $g(x) := f(x) - x$  and apply Newton's method to  $g(x) = 0$ :



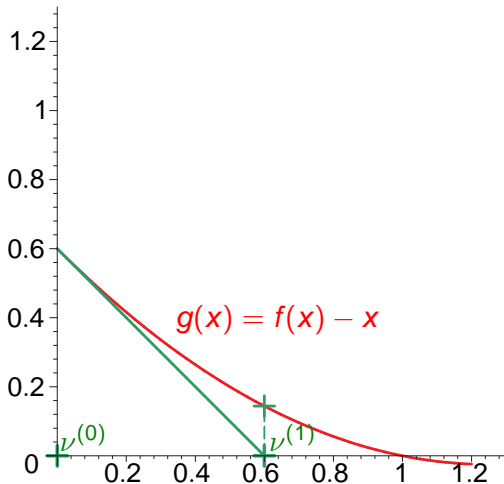
# Newton's Method

Let  $g(x) := f(x) - x$  and apply Newton's method to  $g(x) = 0$ :



# Newton's Method

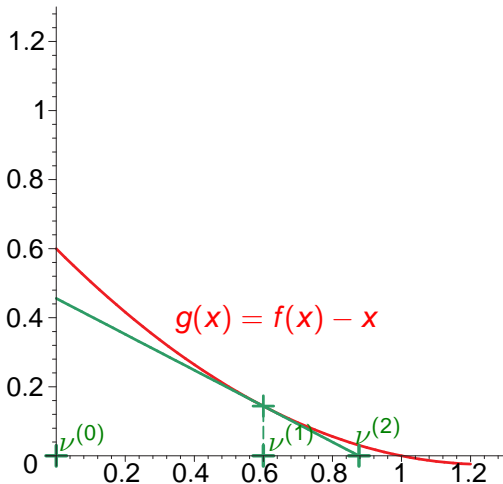
Let  $g(x) := f(x) - x$  and apply Newton's method to  $g(x) = 0$ :





# Newton's Method

Let  $g(x) := f(x) - x$  and apply Newton's method to  $g(x) = 0$ :

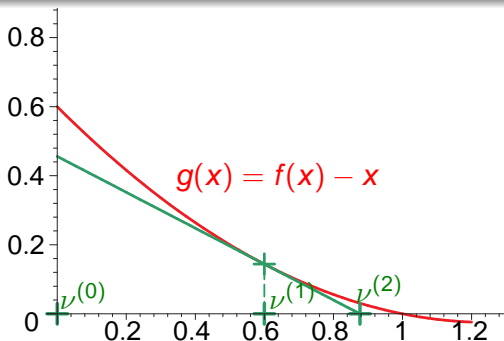


# Newton's Method

Let  $g(x) := f(x) - x$  and apply Newton's method to  $g(x) = 0$ :

Proposition (Etessami, Yannakakis, 2005)

*Newton's method converges to the least solution.*



# Newton's Method

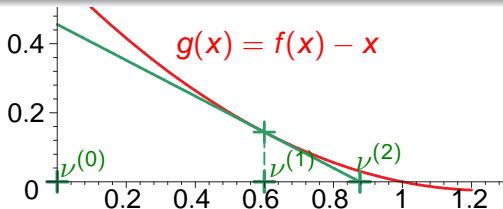
Let  $g(x) := f(x) - x$  and apply Newton's method to  $g(x) = 0$ :

Proposition (Eteessami, Yannakakis, 2005)

*Newton's method converges to the least solution.*

The least solution is  $= 1$ .

Why is he talking about Newton's method??



# Newton's Method

Let  $g(x) := f(x) - x$  and apply Newton's method to  $g(x) = 0$ :

1.2

Proposition (Etessami, Yannakakis, 2005)

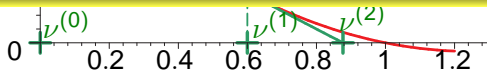
*Newton's method converges to the least solution.*

The least solution is  $= 1$ .

Why is he talking about Newton's method??

Theorem

$$\Pr \left[ S^{op} \leq k \right] = \nu^{(k)}$$



# Tail Bounds for the Optimal Scheduler

## Theorem

$$\Pr \left[ S^{op} \leq k \right] = \nu^{(k)}$$

It follows:

$$\Pr \left[ S^{op} \geq k \right] = 1 - \nu^{(k-1)}$$

The  $\nu^{(k)}$  converge to 1, so  $\Pr \left[ S^{op} \geq k \right]$  goes to 0.

But how fast??

# Tail Bounds for the Optimal Scheduler

## Theorem

$$\Pr \left[ S^{op} \leq k \right] = \nu^{(k)}$$

It follows:

$$\Pr \left[ S^{op} \geq k \right] = 1 - \nu^{(k-1)}$$

The  $\nu^{(k)}$  converge to 1, so  $\Pr \left[ S^{op} \geq k \right]$  goes to 0.

But how fast??

## Corollary (follows from KLE'07, EKL'08)

*general task systems:*  $\Pr \left[ S^{op} \geq k \right] \in \mathcal{O}(d^k) \quad (d < 1)$

*subcritical task systems:*  $\Pr \left[ S^{op} \geq k \right] \in \mathcal{O}(d^{2^k}) \quad (d < 1)$

X

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

X

# Online Scheduling

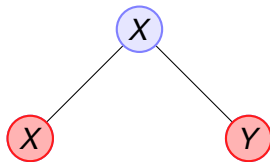
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$



$$X \Rightarrow XY$$



# Online Scheduling

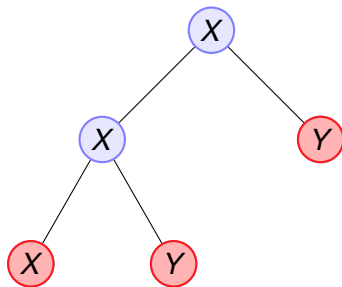
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$



$$X \Rightarrow XY \Rightarrow \color{red}{XYY}$$

# Online Scheduling

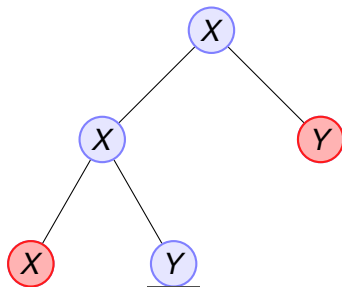
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$



$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY$$

# Online Scheduling

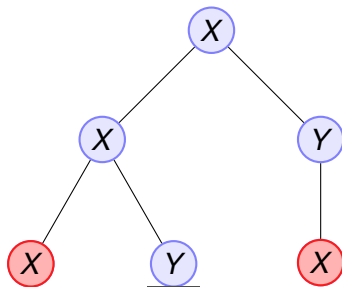
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$



$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow \color{red}{XX}$

# Online Scheduling

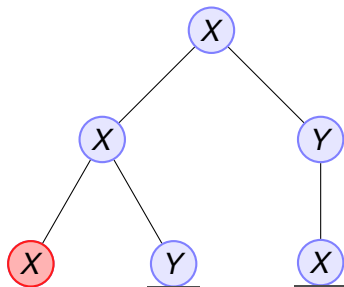
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$



$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X$

# Online Scheduling

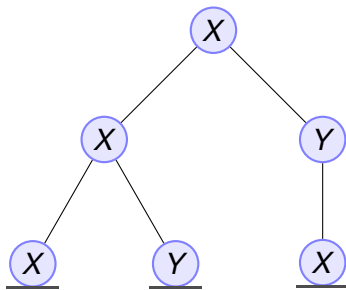
$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$



$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$



The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow \color{red}{XY}Y \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow \color{red}{XY}Y \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow \color{red}{XX} \Rightarrow X \Rightarrow \varepsilon$$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow \color{red}{XX} \Rightarrow X \Rightarrow \varepsilon$$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$



The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$

The **urn model** is more appropriate:

$$X \xrightarrow{0.2} XY$$

$$X \xrightarrow{0.8} \varepsilon$$

$$Y \xrightarrow{0.3} YY$$

$$Y \xrightarrow{0.1} X$$

$$Y \xrightarrow{0.6} \varepsilon$$

$$X \Rightarrow XY \Rightarrow XYY \Rightarrow XY \Rightarrow XX \Rightarrow X \Rightarrow \varepsilon$$

# Weights: An Auxiliary Notion

Let  $\mathbf{v} > \mathbf{1}$  be a vector with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$ .

Choose  $h > 1$  and for all types  $X$  a **weight**  $w_X$  with

$$h^{w_X} = \mathbf{v}_X \quad \text{for all types } X.$$



Denote by  $W$  the **maximum weight** of a derivation.

For instance:  $X \Rightarrow XY \Rightarrow Y \Rightarrow YY \Rightarrow Y \Rightarrow \varepsilon$  yields

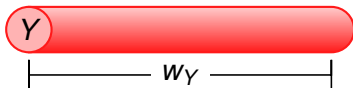
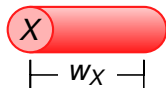


# Weights: An Auxiliary Notion

Let  $\mathbf{v} > \mathbf{1}$  be a vector with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$ .

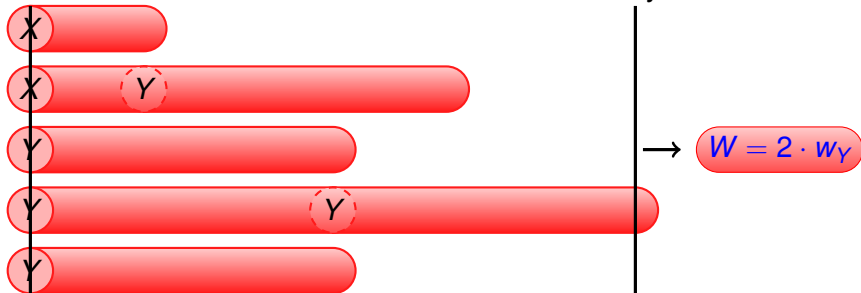
Choose  $h > 1$  and for all types  $X$  a **weight**  $w_X$  with

$$h^{w_X} = \mathbf{v}_X \quad \text{for all types } X.$$



Denote by  $W$  the **maximum weight** of a derivation.

For instance:  $X \Rightarrow XY \Rightarrow Y \Rightarrow YY \Rightarrow Y \Rightarrow \varepsilon$  yields



# An Upper Bound for All Online Schedulers

(Recall:  $\mathbf{v} > \mathbf{1}$  with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$  and  $h^{w_x} = \mathbf{v}_x$ .)

One can show by a martingale argument:

$$\Pr [ W \geq k ] \leq \frac{\mathbf{v}_{X_0}}{h^k}$$

Note: Whenever  $S \geq k$  then  $W \geq k \cdot w_{min}$ .

So we obtain:

$$\Pr [ S \geq k ] \leq \Pr [ W \geq k \cdot w_{min} ] \leq \frac{\mathbf{v}_{X_0}}{h^{k \cdot w_{min}}} = \frac{\mathbf{v}_{X_0}}{\mathbf{v}_{min}^k}$$

# An Upper Bound for All Online Schedulers

(Recall:  $\mathbf{v} > \mathbf{1}$  with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$  and  $h^{w_{X_0}} = \mathbf{v}_{X_0}$ .)

One can show by a martingale argument:

$$\Pr [ W \geq k ] \leq \frac{\mathbf{v}_{X_0}}{h^k}$$

Note: Whenever  $S \geq k$  then  $W \geq k \cdot w_{min}$ .  
So we obtain:

$$\Pr [ S \geq k ] \leq \Pr [ W \geq k \cdot w_{min} ] \leq \frac{\mathbf{v}_{X_0}}{h^{k \cdot w_{min}}} = \frac{\mathbf{v}_{X_0}}{\mathbf{v}_{min}^k}$$

# An Upper Bound for All Online Schedulers

(Recall:  $\mathbf{v} > \mathbf{1}$  with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$  and  $h^{w_X} = \mathbf{v}_X$ .)

One can show by a martingale argument:

$$\Pr [ W \geq k ] \leq \frac{\mathbf{v}_{X_0}}{h^k}$$

Note: Whenever  $S \geq k$  then  $W \geq k \cdot w_{\min}$ .  
So we obtain:

$$\Pr [ S \geq k ] \leq \Pr [ W \geq k \cdot w_{\min} ] \leq \frac{\mathbf{v}_{X_0}}{h^{k \cdot w_{\min}}} = \frac{\mathbf{v}_{X_0}}{\mathbf{v}_{\min}^k}$$

## Theorem

Let  $\mathbf{v} > \mathbf{1}$  with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$ . Then

$$\Pr [ S^\sigma \geq k ] \leq \frac{\mathbf{v}_{X_0}}{\mathbf{v}_{\min}^k}$$

for all online schedulers  $\sigma$  and all  $k \in \mathbb{N}$ .

# An Upper Bound for All Online Schedulers

(Recall:  $\mathbf{v} > \mathbf{1}$  with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$  and  $h^{w_X} = \mathbf{v}_X$ .)

One can show by a martingale argument:

$$\Pr [ W \geq k ] \leq \frac{\mathbf{v}_{X_0}}{h^k}$$

Note: Whenever  $S \geq k$  then  $W \geq k \cdot w_{\min}$ .  
So we obtain:

$$\Pr [ S \geq k ] \leq \Pr [ W \geq k \cdot w_{\min} ] \leq \frac{\mathbf{v}_{X_0}}{h^{k \cdot w_{\min}}} = \frac{\mathbf{v}_{X_0}}{\mathbf{v}_{\min}^k}$$

## Theorem

Let  $\mathbf{v} > \mathbf{1}$  with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$ . Let  $\mathbf{u} > \mathbf{1}$  with  $\mathbf{u} \leq \mathbf{f}(\mathbf{u})$ . Then

$$\frac{c}{\mathbf{u}_{\max}^k} \leq \Pr [ S^\sigma \geq k ] \leq \frac{\mathbf{v}_{X_0}}{\mathbf{v}_{\min}^k}$$

for all online schedulers  $\sigma$  and all  $k \in \mathbb{N}$ .



# An Upper Bound for All Online Schedulers

## Example

Consider the task system from the beginning:

$$f_x(x, y) = 0.2xy + 0.8$$

$$f_y(x, y) = 0.3y^2 + 0.1x + 0.6$$

One can show:  $\mathbf{f}$  has two **fixed points**:  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  and  $\begin{pmatrix} 1.4 \\ 2.2 \end{pmatrix}$ .

So:  $\frac{c}{2.2^k} \leq \Pr \left[ S^\sigma \geq k \right] \leq \frac{1.4}{1.4^k}$  holds for all  $\sigma$ .

## Theorem

Let  $\mathbf{v} > \mathbf{1}$  with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$ . Let  $\mathbf{u} > \mathbf{1}$  with  $\mathbf{u} \leq \mathbf{f}(\mathbf{u})$ . Then

$$\frac{c}{\mathbf{u}_{\max}^k} \leq \Pr \left[ S^\sigma \geq k \right] \leq \frac{\mathbf{v}_{X_0}}{\mathbf{v}_{\min}^k}$$

for *all online schedulers*  $\sigma$  and all  $k \in \mathbb{N}$ .

# An Upper Bound for All Online Schedulers

## Example

Consider the task system from the beginning:

$$f_x(x, y) = 0.2xy + 0.8$$

$$f_y(x, y) = 0.3y^2 + 0.1x + 0.6$$

One can show:  $f$  has two **fixed points**:  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  and  $\begin{pmatrix} 1.4 \\ 2.2 \end{pmatrix}$ .

So:  $\frac{c}{2.2^k} \leq \Pr \left[ S^\sigma \geq k \right] \leq \frac{1.4}{1.4^k}$  holds for all  $\sigma$ .

## Theorem

Let  $\mathbf{v} > \mathbf{1}$  with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$ . Let  $\mathbf{u} > \mathbf{1}$  with  $\mathbf{u} \leq \mathbf{f}(\mathbf{u})$ . Then

$$\frac{c}{\mathbf{u}_{\max}^k} \leq \Pr \left[ S^\sigma \geq k \right] \leq \frac{\mathbf{v}_{X_0}}{\mathbf{v}_{\min}^k}$$

for *all online schedulers*  $\sigma$  and all  $k \in \mathbb{N}$ .

# A Light-First Scheduler For Our Example

For the upper bound we used:

Whenever  $S \geq k$  then  $W \geq k \cdot w_{min}$ .

We have  $\mathbf{v} = \begin{pmatrix} 1.4 \\ 2.2 \end{pmatrix}$ , so  $w_{min} = w_X$ .

We say,  $X$  is the **lightest** type.

$X \xrightarrow{0.2} XY$

$X \xrightarrow{0.8} \varepsilon$

$Y \xrightarrow{0.3} YY$

$Y \xrightarrow{0.1} X$

$Y \xrightarrow{0.6} \varepsilon$

**Light-First Scheduler:**

Process the lightest type (here:  $X$ )  
whenever it is in the pool.

In our example, the light-first scheduler guarantees:  
at any time **at most one**  $X$ -task in the pool.

Hence, with the light-first scheduler:

Whenever  $S \geq k$  then  $W \geq 1 \cdot w_X + (k-1) \cdot w_Y$ .

# A Light-First Scheduler For Our Example

For the upper bound we used:

Whenever  $S \geq k$  then  $W \geq k \cdot w_X$ .

We have  $\mathbf{v} = \begin{pmatrix} 1.4 \\ 2.2 \end{pmatrix}$ , so  $w_{min} = w_X$ .

We say,  $X$  is the **lightest** type.

$X \xrightarrow{0.2} XY$

$X \xrightarrow{0.8} \varepsilon$

$Y \xrightarrow{0.3} YY$

$Y \xrightarrow{0.1} X$

$Y \xrightarrow{0.6} \varepsilon$

**Light-First Scheduler:**

Process the lightest type (here:  $X$ )  
whenever it is in the pool.

In our example, the light-first scheduler guarantees:  
at any time **at most one**  $X$ -task in the pool.

Hence, with the light-first scheduler:

Whenever  $S \geq k$  then  $W \geq 1 \cdot w_X + (k-1) \cdot w_Y$ .

# A Light-First Scheduler For Our Example

Previously, for **all online schedulers**  $\sigma$ :

$$\frac{c}{2.2^k} \leq \Pr \left[ S^\sigma \geq k \right] \leq \frac{1.4}{1.4^k}$$

Now, for the **light-first scheduler**  $lf$ :

$$\Pr \left[ S^{lf} \geq k \right] \leq \frac{1.4}{1.4 \cdot 2.2^{k-1}}$$

So,  $lf$  achieves the **optimal** tail bound:

$$\Pr \left[ S^{lf} \geq k \right] \in \Theta \left( \frac{1}{2.2^k} \right)$$

# A Light-First Scheduler For Our Example

Previously, for **all online schedulers**  $\sigma$ :

$$\frac{c}{2.2^k} \leq \Pr [ S^\sigma \geq k ] \leq \frac{1.4}{1.4^k}$$

$$W \geq k \cdot w_X$$

Now, for the **light-first scheduler**  $lf$ :

$$\Pr [ S^{lf} \geq k ] \leq \frac{1.4}{1.4 \cdot 2.2^{k-1}}$$

So,  $lf$  achieves the **optimal** tail bound:

$$\Pr [ S^{lf} \geq k ] \in \Theta \left( \frac{1}{2.2^k} \right)$$

# A Light-First Scheduler For Our Example

Previously, for **all online schedulers**  $\sigma$ :

$$\frac{c}{2.2^k} \leq \Pr [ S^\sigma \geq k ] \leq \frac{1.4}{1.4^k}$$

$$W \geq k \cdot w_X$$

Now, for the **light-first scheduler**  $lf$ :

$$W \geq 1 \cdot w_X + (k-1) \cdot w_Y$$

$$\Pr [ S^{lf} \geq k ] \leq \frac{1.4}{1.4 \cdot 2.2^{k-1}}$$

So,  $lf$  achieves the **optimal** tail bound:

$$\Pr [ S^{lf} \geq k ] \in \Theta \left( \frac{1}{2.2^k} \right)$$

# Light-First Schedulers in General

More general:

Light-First Scheduler:

Process the lightest type available in the pool.

This strategy improves the general bound on  $\Pr [ S \geq k ]$ .



# Light-First Schedulers in General

More general:

Light-First Scheduler:

Process the lightest type available in the pool.

This strategy improves the general bound on  $\Pr [ S \geq k ]$ .

Q: What is a light type?

A: A type whose component in the vector  $\mathbf{v}$  is small.

with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$

# Light-First Schedulers in General

More general:

Light-First Scheduler:

Process the lightest type available in the pool.

This strategy improves the general bound on  $\Pr [ S \geq k ]$ .

Q: What is a light type?

A: A type whose component in the vector  $\mathbf{v}$  is small.

with  $\mathbf{v} \geq \mathbf{f}(\mathbf{v})$

Q: Well, what is a light type, intuitively?

A: Depends on  $\mathbf{v}$ , but if we compute  $\mathbf{v}$  in a simple way, then the lightest type is the one with the smallest expected **time** (!)

# Comparison

Notation:  $op$  = the optimal offline scheduler  
 $\sigma$  = any online scheduler  
 $0 < d < 1$

In the **subcritical** case:

	tail bound	Expectation
$S^{op}$	$\Pr [ S^{op} \geq k ] \in \mathcal{O}(d^{2k})$	finite
$S^\sigma$	$\Pr [ S^\sigma \geq k ] \in \Theta(d^k)$	finite

# Comparison

Notation:  $op$  = the optimal offline scheduler  
 $\sigma$  = any online scheduler  
 $0 < d < 1$

In the **subcritical** case:

	tail bound	Expectation
$S^{op}$	$\Pr [ S^{op} \geq k ] \in \mathcal{O}(d^{2k})$	finite
$S^\sigma$	$\Pr [ S^\sigma \geq k ] \in \Theta(d^k)$	finite

In the **critical** case:

	tail bound	Expectation
$S^{op}$	$\Pr [ S^{op} \geq k ] \in \mathcal{O}(d^k)$	finite
$S^\sigma$	$\Pr [ S^\sigma \geq k ] \in \Omega\left(\frac{1}{k}\right)$	infinite

# Conclusions

- We have studied the **space consumption** of stochastically generated tasks.
- Such task systems require **scheduling**.
- The performance of the **optimal offline scheduler** is closely linked with the convergence speed of **Newton's method**.
- The performance of the **online schedulers** is closely linked with **fixed points of the function  $f$** .
- **Light-First** schedulers are **good** online schedulers.
- For critical systems, **finite expectation** is only achieved by offline schedulers.
- One can **efficiently approximate expectations**.

Thank you!