Minimax Games with Bandits
Open problem for Colt09

Jake & Manfred
UCB & UCSC

June 19, 2009

Expert setting

loop
  produce probility vector $\mathbf{w}$ over n experts
  receive loss vector $\boldsymbol{\ell} \in \{0, 1\}^n$
  incur loss $\mathbf{w} \cdot \boldsymbol{\ell}$
end loop

We have a problem

	full information	bandit
multiplicative	WMR/Hedge	Exp3
minimax	last year	??your turn??

Updates

	full information	bandit
multiplicative	$w_i \sim e^{-\eta \, \mathrm{loss}_i}$	estimate losses
minimax	random playout	??

Minimax formulation

- State $\mathbf{s}$ = total loss vector for the expert

$$V(\mathbf{s}) \stackrel{?}{=} \min_{\text{dist. } \mathbf{w}} \max_{\ell \in \{0,1\}^n} \mathbf{w} \cdot \ell + V(\mathbf{s} + \ell)$$

- Assumption: loss of best is $\leq k$

Minimax formulation

- State $\mathbf{s}$ = total loss vector for the expert

$$V(\mathbf{s}) \stackrel{?}{=} \min_{\text{dist. } \mathbf{w}} \max_{\boldsymbol{\ell} \in \{0,1\}^n} \mathbf{w} \cdot \boldsymbol{\ell} + V(\mathbf{s} + \boldsymbol{\ell})$$

- Assumption: loss of best is $\leq k$

Minimax formulation

- State $\mathbf{s}$ = total loss vector for the expert

$$V(\mathbf{s}) \stackrel{?}{=} \min_{\text{dist. } \mathbf{w}} \max_{\boldsymbol{\ell} \in \{0,1\}^n} \mathbf{w} \cdot \boldsymbol{\ell} + V(\mathbf{s} + \boldsymbol{\ell})$$

- Assumption: loss of best is $\leq k$

Minimax formulation refined

Dead state $\mathbf{d} = (k+1, k+1, \ldots, k+1)$

$$V(\mathbf{d}) = 0$$

Otherwise

$$V(\mathbf{s}) = \min_{\text{dist. } \mathbf{w}} \max_{\ell \in \{0,1\}^n} \mathbf{w} \cdot \boldsymbol{\ell} + V(\mathbf{s} \dot{+} \boldsymbol{\ell})$$

The optimum probability of expert i at state $\mathbf{s}$ is
the probability that this expert is the last surviving expert
when the loss is assigned randomly

Minimax formulation refined

Dead state $\mathbf{d} = (k+1, k+1, \ldots, k+1)$

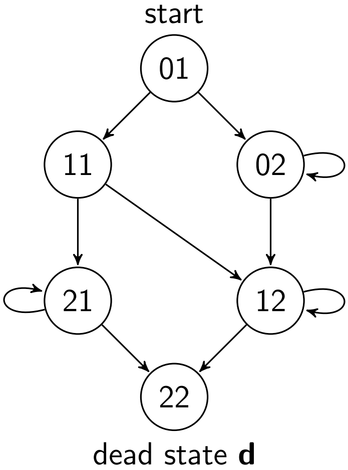
$$V(\mathbf{d}) = 0$$

Otherwise

$$V(\mathbf{s}) = \min_{\text{dist. } \mathbf{w}} \max_{\boldsymbol{\ell} \in \{0,1\}^n} \mathbf{w} \cdot \boldsymbol{\ell} + V(\mathbf{s} \dot{+} \boldsymbol{\ell})$$

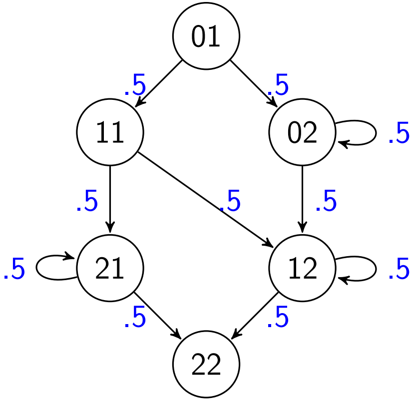
The optimum probability of expert i at state $\mathbf{s}$ is
the probability that this expert is the last surviving expert
when the loss is assigned randomly

Visualizing the optimal algorithm



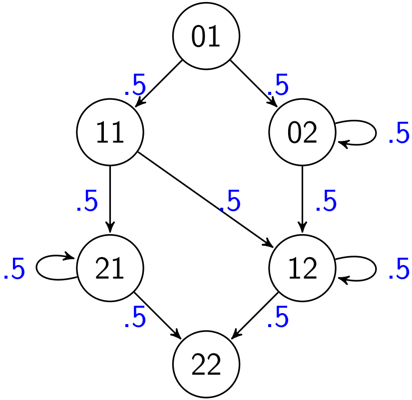
$k = 1$

Random paths



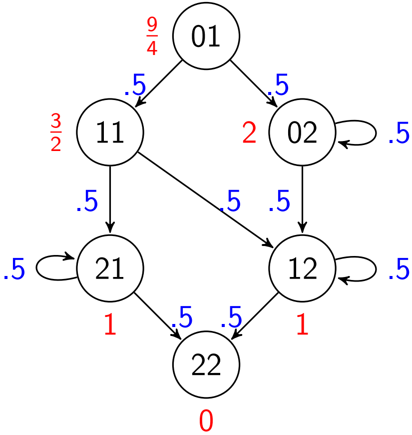
Value = $\frac{\text{expected pathlength}}{n}$

Random paths



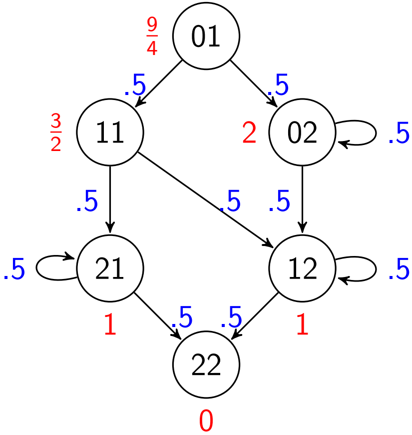
$\text{Value} = \frac{\text{expected pathlength}}{n}$

Values



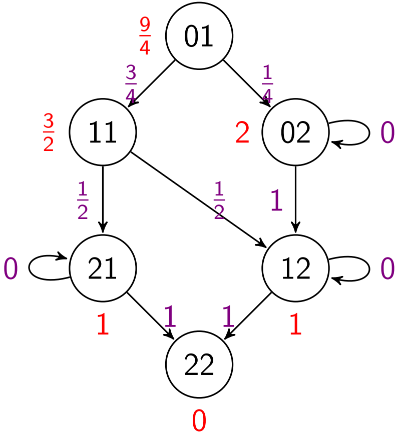
Optimum probabilities = gap of values

Values



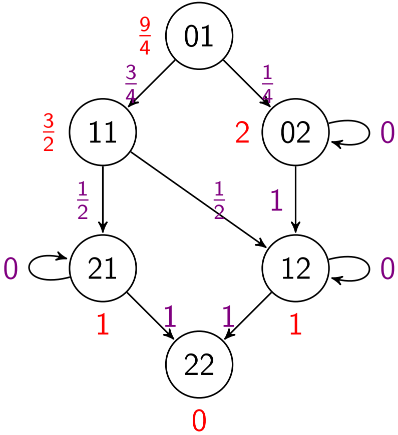
Optimum probabilities = gap of values

Optimum probabilities



- Optimum probability i = probability of paths with i last
- Use random path to predict

Optimum probabilities



- Optimum probability i = probability of paths with i last
- Use random path to predict

Back to your job

Minimax algorithm for the bandit case

Some help?

- What is a natural "sufficient statistic" of the state of the game?
- What is a natural "stopping criterion" for the game?
- In the hedge setting game, the optimal strategy can be characterized using the notion of a "random play-out".

  Does the minimax strategy in the bandit setting admit a similar characterization?

Back to your job

Minimax algorithm for the bandit case

Some help?

- What is a natural "sufficient statistic" of the state of the game?
- What is a natural "stopping criterion" for the game?
- In the hedge setting game, the optimal strategy can be characterized using the notion of a "random play-out".

  Does the minimax strategy in the bandit setting admit a similar characterization?

Back to your job

Minimax algorithm for the bandit case

Some help?

- What is a natural "sufficient statistic" of the state of the game?
- What is a natural "stopping criterion" for the game?
- In the hedge setting game, the optimal strategy can be characterized using the notion of a "random play-out".

  Does the minimax strategy in the bandit setting admit a similar characterization?

Back to your job

Minimax algorithm for the bandit case

Some help?

- What is a natural “sufficient statistic” of the state of the game?
- What is a natural “stopping criterion” for the game?
- In the hedge setting game, the optimal strategy can be characterized using the notion of a “random play-out”.

  Does the minimax strategy in the bandit setting admit a similar characterization?

Back to your job

Minimax algorithm for the bandit case

Some help?

- What is a natural "sufficient statistic" of the state of the game?
- What is a natural "stopping criterion" for the game?
- In the hedge setting game, the optimal strategy can be characterized using the notion of a "random play-out".

  Does the minimax strategy in the bandit setting admit a similar characterization?

$ 200