Combinatorics of Optimal Stopping on Partial Orders and Graphs

Michał Morayne

Wroclaw University of Technology

2nd Polish Combinatorial Conference

Banach Center 2008

The classical secretary problem

In the classical secretary problem there are n linearly ordered elements:
They are being observed at some random order $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$. At the moment $t = i$ the observer knows only the relative ranks of the elements examined so far.
The aim of the observer is to choose the currently examined object in such a way that the probability

$$P[\omega_\tau = n]$$

is the maximal possible.

Whether $\tau(\omega) = i$ depends only on

$$(\omega_1, \omega_2, \ldots, \omega_i).$$

Such a random variable τ is called a *stopping time.*

The *optimal stopping time* $\tau^{(n)}$, i.e. such that

$$P[\omega_{\tau^{(n)}} = n] = \max_{\tau} P[\omega_\tau = n]$$

in the classical (linear) case is equal to

$$\tau^{(n)} = \min\{i : i \geq M_n \text{ and } \omega_i > \omega_1, \ldots, \omega_{i-1}\},$$

where

$$M_n = \min\{m : \sum_{j=m}^{n-1} \frac{1}{j} \leq 1\}.$$

We also have

$$P[\omega_{\tau^{(n)}} = n] \to \frac{1}{e},$$

and

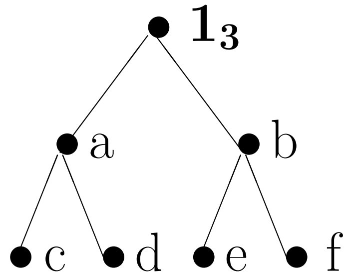
$$\frac{M_n}{n} \to \frac{1}{e},$$

with $n \to \infty$.

Best choice problem for partially ordered sets

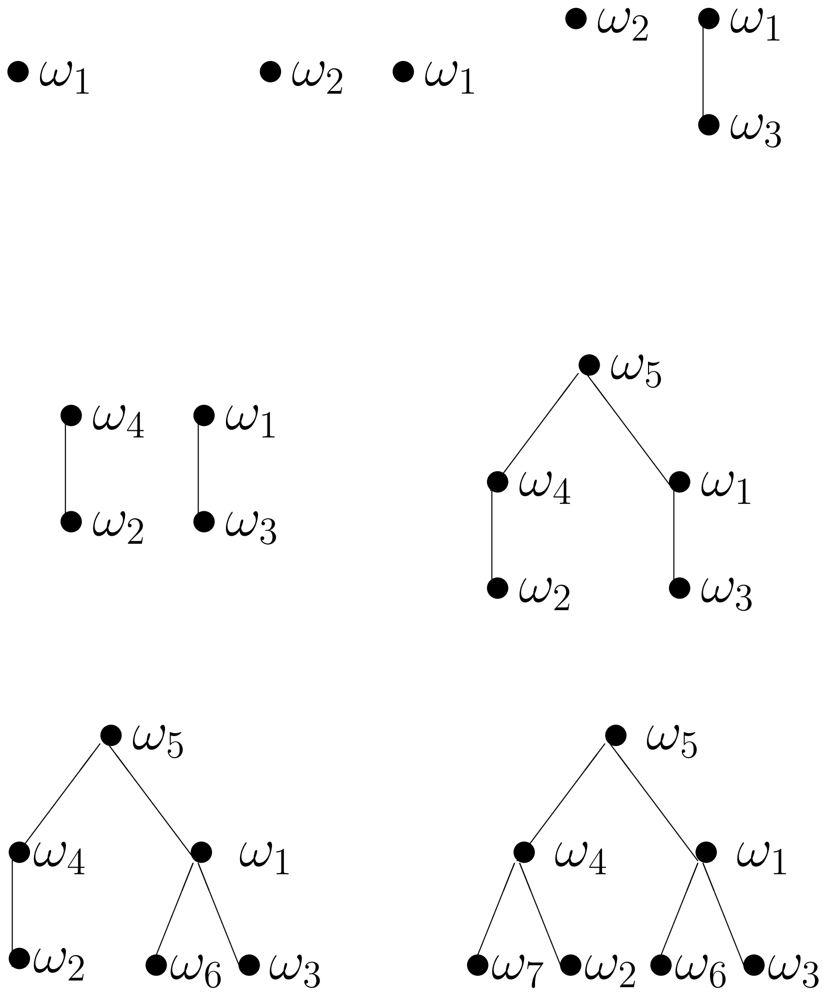
One can extend the secretary problem to a partial order version.

Let us consider the example of $\mathbf{T_n}$, i.e. the complete binary tree of depth n, for $n = 3$.



Assume that the observer deals with the following sequence of objects

$$(b, d, f, a, \mathbf{1_3}, e, c).$$



ω_1
ω_2 ω_1
ω_2 ω_1 ω_3
ω_4 ω_1 ω_2 ω_3
ω_5 ω_4 ω_1 ω_2 ω_3
ω_5 ω_4 ω_1 ω_2 ω_6 ω_3
ω_5 ω_4 ω_1 ω_7 ω_2 ω_6 ω_3

In the case of the complete binary tree $(\mathbf{T_n}, \prec)$ the optimal stopping time $\tau^{(n)}(\omega)$ is the minimal i such that

1) $\omega_1, \ldots, \omega_{i-1} \prec \omega_i$

and

2) either
$\{\omega_1, \ldots, \omega_i\}$ is not a chain
or
$\{\omega_1, \ldots, \omega_i\}$ is a chain and $i > n/2$.

[M. Morayne, Partial order version of the secretary problem, the binary tree case, Discrete Mathematics 184 (1998), 165-181].

The combinatorial fact that implies this theorem is stated in the theorem below.

Let T be a finite tree. Let $\mathbf{1}_n$ be the root of $\mathbf{T}_n$. Let

$$A_T^{(n)} = |\{A \subseteq \mathbf{T_n} : A \cong T \text{ and } \mathbf{1_n} \in A\}|$$

and

$$B_T^{(n)} = |\{A \subseteq \mathbf{T_n} : A \cong T \text{ and } \mathbf{1_n} \notin A\}|.$$

Theorem. *If T is a nonlinear tree, then*

$$A_T^{(n)} > B_T^{(n)},$$

or, equivalently,

$$\frac{A_T^{(n)}}{B_T^{(n)}} > 1, \quad \left(\frac{A_T^{(n)}}{A_T^{(n)} + B_T^{(n)}} > \frac{1}{2}\right).$$

If T is a chain, then

$$A_T^{(n)} > B_T^{(n)}, \quad if \quad |T| > \frac{n}{2}$$

and

$$A_T^{(n)} < B_T^{(n)}, \quad if \quad |T| \leq \frac{n}{2}.$$

Note that

$$P[\omega_i = \mathbf{1_n} | \{\omega_1, \ldots, \omega_i\} \cong T \\ \text{and } \ \omega_i = \max\{\omega_1 \ldots, \omega_i\}] = \\ \frac{A_T^{(n)}}{A_T^{(n)} + B_T^{(n)}}.$$

If we perform two independent searches on $\mathbf{T_n}$ according to the optimal strategy $\tau^{(n)}$ and the first ends at the structure

$$\{\omega'_1, \ldots, \omega'_{\tau^{(n)}(\omega')}\} \cong T_1$$

and the second at a structure

$$\{\omega''_1, \ldots, \omega''_{\tau^{(n)}(\omega'')}\} \cong T_2$$

and T_1 is isomorphic to a subposet of T_2 (we write $T_1 \subset T_2$), then we want to compare the numbers

$$p_1 = P[\omega_{\tau(n)} = \mathbf{1_n} | \{\omega_1, \ldots, \omega_{\tau(n)}\} \cong T_1] = \frac{A_{T_1}^{(n)}}{A_{T_1}^{(n)} + B_{T_1}^{(n)}}$$

and

$$p_2 = P[\omega_{\tau(n)} = \mathbf{1_n} | \{\omega_1, \ldots, \omega_{\tau(n)}\} \cong T_2] = \frac{A_{T_2}^{(n)}}{A_{T_2}^{(n)} + B_{T_2}^{(n)}}.$$

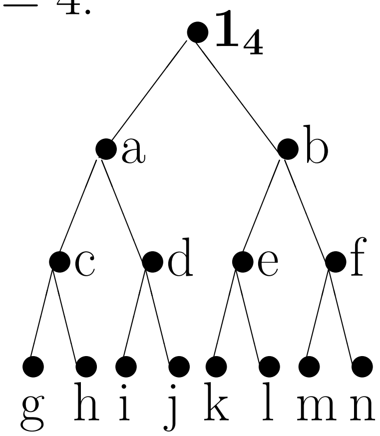
Intuitively,

$$p_1 \leq p_2,$$

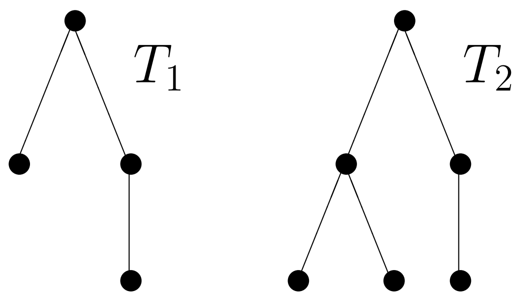
which is equivalent to

$$\frac{A_{T_1}^{(n)}}{B_{T_1}^{(n)}} < \frac{A_{T_2}^{(n)}}{B_{T_2}^{(n)}}.$$

Assume $n = 4$.



and that the first search ended at the tree T_1 and the second at the tree T_2:



For example:

$$\tau_4(c, d, j, \mathbf{1_4}, \ldots) = \mathbf{1_4} \text{ and } \tau_4(k, e, f, b, \ldots) = b,$$

$$\tau_4(c, d, b, a, k, \mathbf{1_4}, \ldots) = \mathbf{1_4}$$

and

$$\tau_4(d, i, j, h, c, a, \ldots) = a.$$

$$\{c, d, j.\mathbf{1_4}\} \cong \{k, e, f, b\} \cong T_1$$

and

$$\{c, d, b, a, k, \mathbf{1_4}\} \cong \{d, i, j, h, c, a\} \cong T_2.$$

Theorem. *If T_1, T_2 are binary trees and $T_1 \subset T_2$, then*

$$\frac{A_{T_1}^{(n)}}{B_{T_1}^{(n)}} \leq \frac{A_{T_2}^{(n)}}{B_{T_2}^{(n)}}.$$

[G. Kubicki, J. Lehel, M. Morayne, A ratio inequality for binary trees and the best secretary, Combinatorics, Probability and Computing 11 (2002), 149-161]

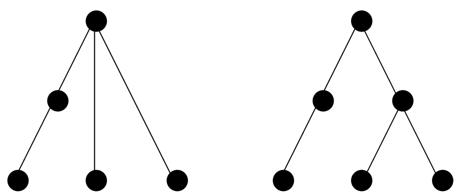
We posed the following:

False Conjecture: that the conclusion was true without the assumption that T_1, T_2 are binary (\$ 100 + dinner).

And the winner is:

Nicholas Georgiou (Bristol, UK).

A counterexample (for $n \neq 6, \ldots, 11$)



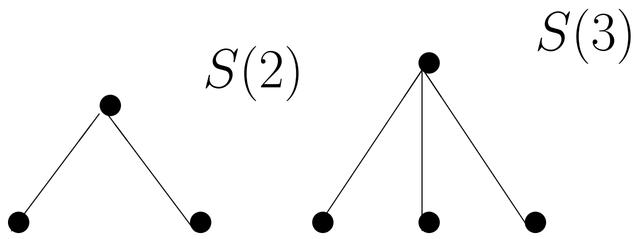
[N. Georgiou, Embeddings and other mappings of rooted trees into complete trees, Order 22 (2005), 257-288]

The theorem is also true when T_1, T_2 are stars.

Let $S(i)$ be the i-star.

Theorem. *If $k \leq m$ then*

$$\frac{A^{(n)}_{S(k)}}{B^{(n)}_{S(k)}} \leq \frac{A^{(n)}_{S(m)}}{B^{(n)}_{S(m)}}.$$



[G. Kubicki, J. Lehel, M. Morayne, Counting chains and antichains in the complete binary tree, Ars Combin. 79 (2006), 245–256.].

The asymptotic version of the above conjecture is true. It follows from the following theorem.

Theorem.

$$\lim_{n\to\infty} \frac{A_T^{(n)}}{B_T^{(n)}} = 2^{l(T)-1} - 1,$$

where $l(T)$ is the number of leaves of T.

Corollary. *If $T_2 \subset T_1$, then*

$$\lim_{n\to\infty} \frac{A_{T_2}^{(n)}}{B_{T_2}^{(n)}} \leq \lim_{n\to\infty} \frac{A_{T_1}^{(n)}}{B_{T_1}^{(n)}}.$$

[G. Kubicki, J. Lehel, M. Morayne, An asymptotic ratio in the complete binary tree, Order 20 (2003), 91–97].

Let $\mathcal{T}$ be now any tree. Let $\mathbf{1}_{\mathcal{T}}$ be the root (the maximal element) of $\mathcal{T}$ and let T be any subtree of $\mathcal{T}$.

$$A_T^{\mathcal{T}} = |\{A \subseteq \mathcal{T} : A \cong T \text{ and } \mathbf{1}_{\mathcal{T}} \in A\}|$$

and

$$B_T^{\mathcal{T}} = |\{A \subseteq \mathcal{T} : A \cong T \text{ and } \mathbf{1}_{\mathcal{T}} \notin A\}|.$$

Looking for optimal stopping times on an arbitrary $\mathcal{T}$ it is also important to know when

$$\frac{A_T^{\mathcal{T}}}{B_T^{\mathcal{T}}} < 1$$

or, equivalently,

$$\frac{A_T^{\mathcal{T}}}{A_T^{\mathcal{T}} + B_T^{\mathcal{T}}} < \frac{1}{2},$$

and how much the last ratio is smaller than $\frac{1}{2}$.

The following two results give two estimations in the situation when $\mathcal{T}$ is arbitrary and T is a chain.

If T is a chain of length k let

$$A_T^{\mathcal{T}} = A(\mathcal{T}, k)$$

and

$$B_T^{\mathcal{T}} = B(\mathcal{T}, k).$$

Let $\mathrm{AD}(\mathcal{T})$ denote the average depths of $\mathcal{T}$, i.e. the arithmetic mean of the depths of all leaves of $\mathcal{T}$.

Theorem. If $\mathcal{T}$ is a tree with $\mathrm{AD}(\mathcal{T}) \geq 2k$ and $\mathcal{T}$ has s leaves of depth $2k$ and at least one leaf of depth $> 2k$ then

$$\frac{A(\mathcal{T}, k))}{A(\mathcal{T}, k) + B(\mathcal{T}, k)} \leq$$

$$\frac{1}{2} - \frac{3}{2} \times \frac{2k+1}{k(s+1)+s+1} \times \frac{k-1}{10k^2+k-3}.$$

[M.Kuchta, MM, J.Niemiec, Counting embeddings of a chain into a tree, Discrete Math. 297 (2005), 49–59]

Theorem. Let $\mathcal{T}$ be a binary tree of depth greater than $2n + n\ln\, n$. Then

$$\frac{A(\mathcal{T},k))}{A(\mathcal{T},k)+B(\mathcal{T},k)} \leq \frac{1}{2}.$$

[M.Kuchta, M. Morayne, J. Niemiec, Counting embeddings of a chain into a binary tree, to appear in Ars Combinatoria].

Preater's Universal Algorithm

For each particular partially ordered set there is a specific stopping time.

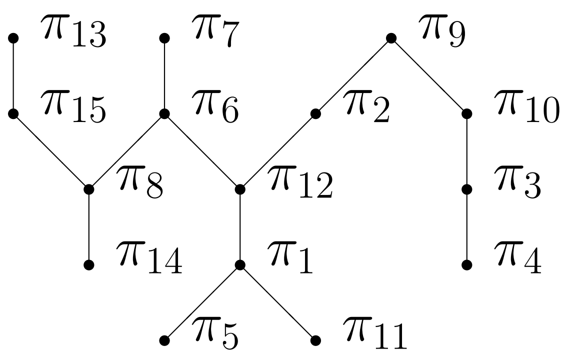
Is there an efficient stopping time defined in a "universal way" for all possible partial orders, where the aim is to get one of the maximal elements.

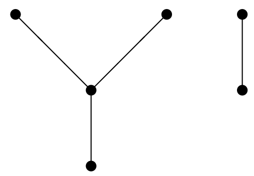
Yes, there is such an algorithm and it is surprisingly efficient! (J. Preater)

Assume that $|X| = n$ is known and $\prec$ is not known to the selector.

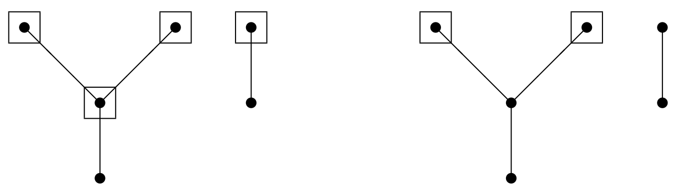
1. Flip a symmetric coin n times. If it comes down tails M times, examine the first M objects of a random permutation of X. They form a set Y. State the height of the set Y, i.e., the maximum length of a chain consisting of elements of Y; denote it by $h(Y)$.
2. Flip a symmetric coin once. If it comes down heads, tag all the objects from Y whose induced height is equal to $h(Y)$; if it comes down tails, tag all the objects from Y whose induced height is greater than or equal to $h(Y) - 1$.
3. Examine further the remaining elements of X. Pick the first object that dominates any of the tagged ones and is maximal up to now. If such an object does not appear, pick the final one.

Call this stopping time: τ.


π_{13}
π_7
π_9
π_{15}
π_6
π_2
π_{10}
π_8
π_{12}
π_3
π_{14}
π_1
π_4
π_5
π_{11}



$M = 6$



Probability of success:

$$P[\pi(\tau) \in MAX] \geq 1/8$$

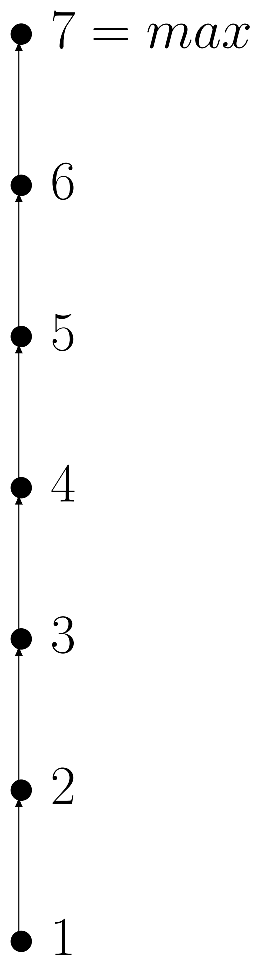
[J. Preater, The best-choice problem for partially ordered objects. Oper. Res. Lett. 25 (1999), 187–190.]

Probability of success:

$$P[\pi(\tau) \in MAX] \geq 1/4$$

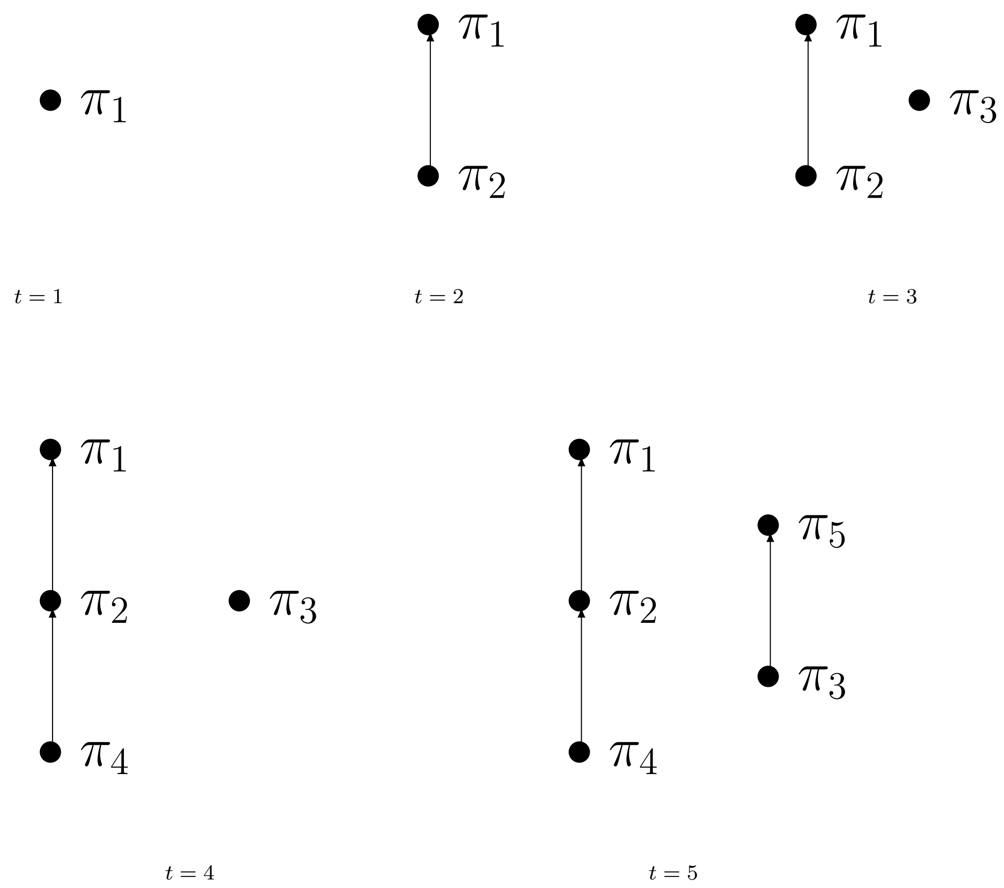
[N. Georgiou, M. Kuchta, M. Morayne, J. Niemiec, On a universal best choice algorithm for partially ordered sets, Random Structures Algorithms 32 (2008), 263–273.]

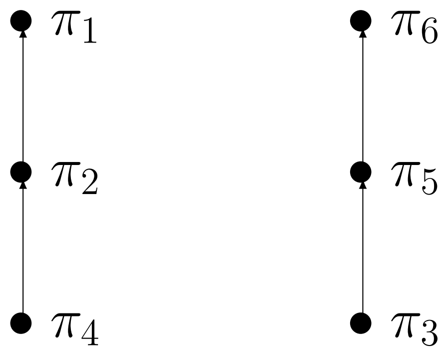
A generalization to the search on graphs; a directed path



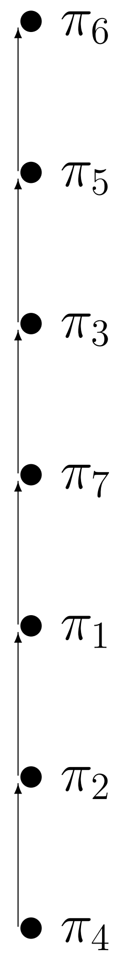
selector's consecutive observations in the case of a directed path of length 7:

Let $\pi = (3, 2, 5, 1, 6, 7, 4)$.





$t = 6$



$t = 7$

Optimal stopping time (τ_n):

Stop when there is a positive conditional (given history) probability that the presently examined candidate is the maximal one and the probability that the maximal one can be among the future candidates is equal to zero.

Probability of success:

$$P[\tau_n(\pi) = max] = \frac{1}{n} \sum_{i=0}^{\lfloor \frac{n-1}{2} \rfloor} \frac{(n-i-1)\ldots(n-2i)}{(n-1)\ldots(n-i)}.$$

Asymptotics:

$$P[P[\tau_n(\pi) = max]\sqrt{n} \to \frac{\sqrt{\pi}}{2}.$$

[G. Kubicki, M. Morayne, Graph-theoretic generalization of the secretary problem: the directed path case. SIAM J. Discrete Math. 19 (2005), 622–632]

Further results:

G. Kubicki and M. Morayne, How to choose the best twins, submitted.

M. Przykucki, M. Sulkowska: Best choice and Gusein-Zade problems for directed and undirected paths, in preparation.

W. Kaźmierczak: Optimal stopping time for a poset whose Hasse diagram is a binary tree with a single branching at the root, in preparation.

Bibliography

W. Stadje, Efficient stopping of a random series of partially ordered points, Proceedings of the III International Conference on Multiple Criteria Decision Making, Köningswinter, 1979, Springer Lecture Notes in Economics and Mathematical Systems, Vol. 177, 1980.

A.V. Gnedin Multicriteria extension of the best choice problem: sequential selection without linear order, Conemp. Math. 125 (1992), 153–172.