A (Gittins) Index Theorem for Randomly Evolving Graphs.

Isaac Sonin (joint with E. Presman)
Dept. of Mathematics, UNC Charlotte, Charlotte NC 28223.

March 1, Computational Finance Seminar in CMU

We consider the problem which informally can be formulated as follows. Initially a finite set of independent trials is available. If a Decision Maker (DM) chooses to test a specific trial she receives a reward depending on the trial tested. As a result of testing, different outcomes can occur. With some probability, which depends on the trial, the process of testing is *terminated* and with complementary probability the tested trial becomes unavailable but some random finite set (possibly empty) of new independent trials is added to the set of available trials. This process can be repeated, but the total number of potential trials is finite. A DM knows the rewards and transition probabilities of all trials. At each step she can either *quit* or continue. Her goal is to select an order to test trials and a quitting time to maximize the expected total reward.

On one side this problem is a generalization of a "least cost testing sequencing" problem solved independently by a few authors in 1960 (see [1]), and on the other side is a generalization of the so called Multi-armed Bandit problem with independent arms (see, for example, [2]). It is also a generalization of [3] where the set of new trials was deterministic. This model allows more general economic interpretation when trials are interpreted as economic projects which bring some rewards but with some probability may also result in the termination of decision process.

Formally the model can be formulated as a Markov Decision Process model defined on a graph which is a directed forest F_0 with a finite number of edges, where each edge represents a possible trial. A state x is either an absorbing state x_*, or consists of a subset (possibly empty) of edges of F_0, available for testing at that moment. For each edge e of the forest, denote by $T(e)$ the directed tree which consists of all edges which follow e. Given an initial state x and strategy π, the goal is to maximize the expected total reward, R_x^π. The classical Gittins model will be a special case of this model if we allow forest F_0 to be infinite and set all termination probabilities equal to $(1-\beta)$ where β is a discount coefficient.

Main Problem: *Given an initial state x, maximize R_x^π over all strategies.*

Let $\gamma(e)$ be a function on a forest F_0. A strategy π is called a (γ, c)-priority rule or simply a priority rule if π tests each time the edge with the highest value of γ among all available edges with values greater than c, and quits otherwise. Let S_e^π be the probability of termination under priority rule π with an initial state $x = e$.

Auxiliary problem: *For any edge e, maximize R_e^π / S_e^π over all priority rules defined on $T(e)$.*

Let $\alpha(e)$ denote the result of maximization in the Auxiliary problem. Our main result is similar to the celebrated Gittins result and the result in [3] but index α has a more general structure.

Theorem. a) *A (unique) optimal strategy in the Main Problem is a $(\alpha, 0)$-priority rule,*

(b) *a (unique) optimal strategy in the Auxiliary Problem is a* $(\alpha, \alpha(e))$*-priority rule defined on* $T(e)$.

References.

[1] Mitten, L. G. (1960) *An Analytic Solution to the Least Cost Testing Sequence Problem. J. of Industr. Eng. 11, no. 1, 17.*

[2] Berry, D. A.; Fristedt B. (1985) *Bandit Problems. Sequential Allocation of Experiments. Monographs on Statistics and Applied Probability, Chapman & Hall, London.*

[3] *Denardo E. V., Rothblum U.G., Van der Heyden L. (2003), Index Policies for Stochastic Search in a Forest with Application to R & D Project Management, manuscript.*