CSE/ECE 848
Introduction to Evolutionary Computation

Module 2, Lecture 8, Part 4b
More Theory—K-Armed Bandit

Erik D. Goodman, Executive Director
BEACON Center for the Study of Evolution in Action
Professor, ECE, ME, and CSE

Exponentially Increasing Sampling and the K-Armed Bandit Problem

- Schema Theorem says instances of a good building block H grow exponentially: M(H,t+1) >= k M(H,t)

  (if we set crossover and mutation rates well)

  That is, H's instances in population grow exponentially, as long as rare relative to pop size and k>1 (H is a "building block").
- Is this a good way to allocate trials to schemata? K-armed Bandit argument says we SHOULD devote exponentially increasing fraction of trials to schemata that have performed better in samples so far…

Two-Armed Bandit Problem (from Goldberg, ‘89)

- 1-armed bandit = slot machine
- 2-armed bandit = slot machine with 2 handles, NOT necessarily yielding same payoff odds (or 2 different slot machines)
- If can make a total of N pulls, how should we proceed, so as to maximize expected final total payoff – Ideas???

Two-Armed Bandit, cont.

- Assume LEFT pays with (unknown to us) expected value m_1 and variance s_1^2, and RIGHT pays m_2, with variance s_2^2.
- The DILEMMA: Must EXPLORE while EXPLOITING. (… Sound like a GA?) Clearly a tradeoff must be made. Given that one arm seems to be paying off better than the other SO FAR, how many trials should be given to the BETTER (so far) arm, and how many to the POORER (so far) arm?

Two-Armed Bandit, cont.

Classical approach: SEPARATE the EXPLORATION from EXPLOITATION: If will do N trials, start by allocating n trials to each arm ($2n<N$) to decide WHICH arm appears to be better, and then allocate ALL remaining ($N-2n$) trials to it.

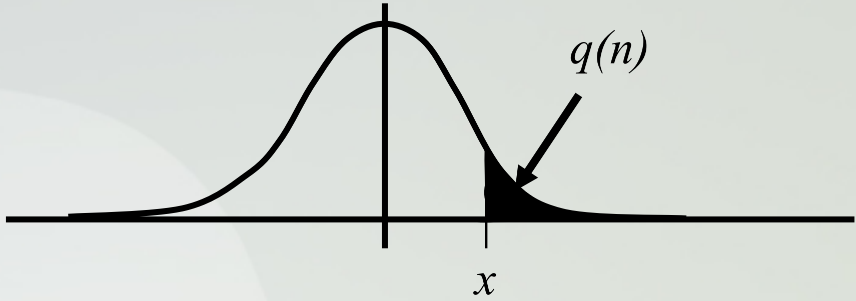
DeJong calculated the expected loss (compared to the OPTIMUM) of using this strategy:

$L(N,n) = |m_1 - m_2| \; . \; [(N-n)\, q(n) + n(1-q(n))]$,where $q(n)$ is the probability that the WORST arm is the OBSERVED BEST arm after n trials on each machine.

Two-Armed Bandit, cont.

This q(n) is well approximated by the tail of the normal distribution:

$$q(n) = \frac{1}{\sqrt{2\pi}} \bullet \frac{e^{-x^2/2}}{x}$$



, where $x = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \bullet \sqrt{n}$

(x is "signal difference to noise ratio" times sqrt(n).)

(Let's call signal difference to noise ratio "c".)

Two-Armed Bandit, cont.

The LARGER x becomes, the LESS probable q(n) becomes (i.e., smaller chance of error). You can see that q(n) (chance of error) DECLINES as n is picked larger, or as the difference in expected values INCREASES or as the sum of the variances DECREASES.

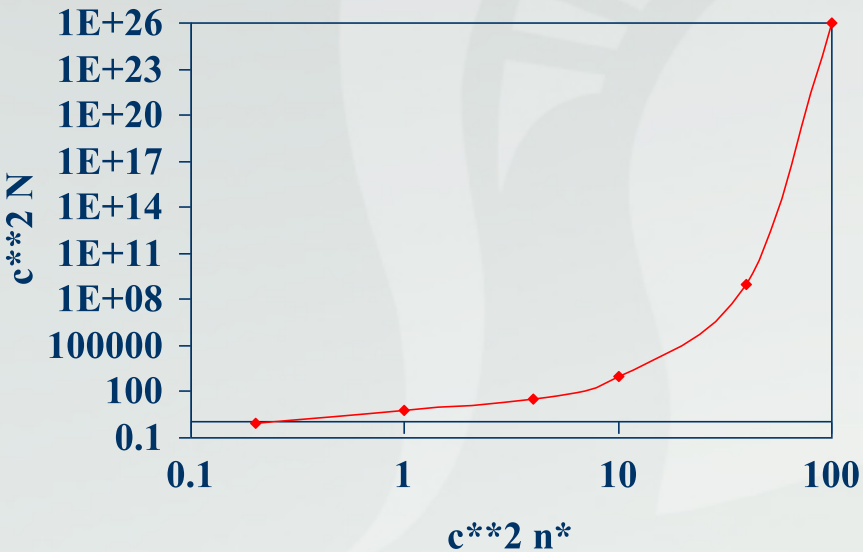
The equation shows two sources of expected loss:

$$L(N,n) = |m_1 - m_2| \,.\, [(N-n)\,q(n) + n(1-q(n))],$$

Due to ^^wrong arm later ^^wrong during exploration

Two-Armed Bandit, cont.

For any N, solve for the optimal experiment size n* by setting the derivative of the loss equation to 0. Graph below (after Fig. 2.2 in Goldberg, ’89) shows the optimal n* as a function of total number of trials, N, and c, the ratio of signal difference to noise.



From graph, see that total number of experiments N grows at a greater-than-exponential function of the ideal number of trials n* in the *exploration* period -- that means, according to classical decision theory, that we should be allocating trials to the BETTER (higher measured “fitness” during the exploration period) of the two arms, at a GREATER THAN EXPONENTIAL RATE.

Two-Armed Bandit, K-Armed Bandit

Now, let our "arms" represent competing schemata. Then the future sampling of the better one (to date) should increase at a larger-than-exponential rate. A GA, using selection, crossover, and mutation, does that (when set properly, according to the schema theorem). If there are K competing schemata over a set of positions, then it is a K-armed bandit.

But at any time, MANY different schemata are being processed, with each competing set representing a K-armed bandit scenario. So maybe the GA's way of allocating trials to schemata is pretty good!