Genetic & Evolutionary Algorithms: Further Investigation on Selection (2)

March 26, 2013

Prof. Chang Wook Ahn

Sungkyunkwan Evolutionary Algorithms Lab.

School of Info. & Comm. Eng.

Sungkyunkwan University

Contents

- **Review**
  - Proportional selection, Scaling, Ranking selection
- **Ordinal Selection**
  - Ranking Selection
  - Tournament Selection
- **Comparison of Selection Pressures**
  - Proportional, Scaling, Ranking, Tournament
- **Takeover Time**
  - Proportional selection & Tournament selection
- **Real-World Applications**
  - Resource Allocation
  - Traveling Salesman Problem, etc.

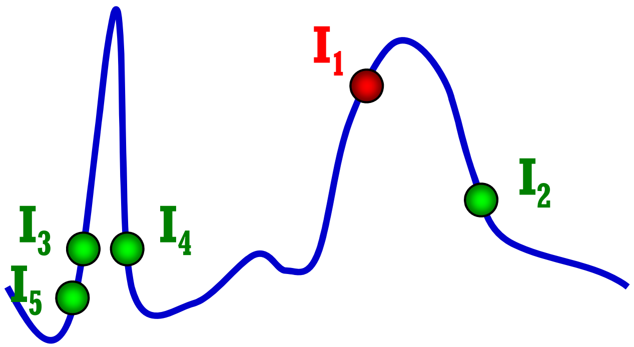
Further Studies on Selection

Review

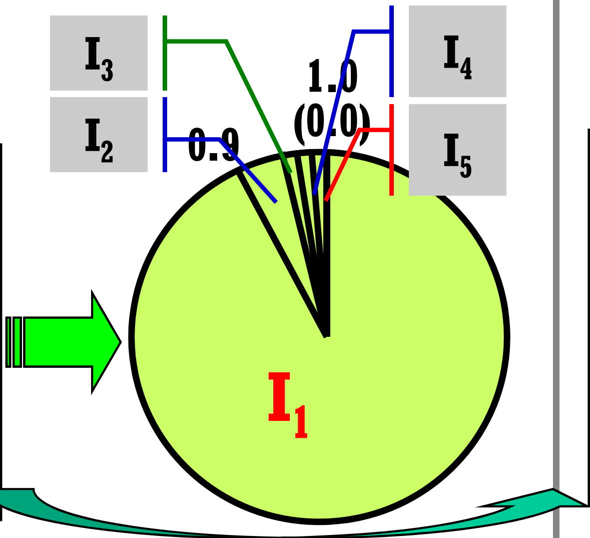
Proportional Selection	Scaling	Ranking Selection

Proportional Selection

- Premature convergence may happen



I_1	90
I_2	5
I_3	2
I_4	2
I_5	1



Fitness Roulette Wheel

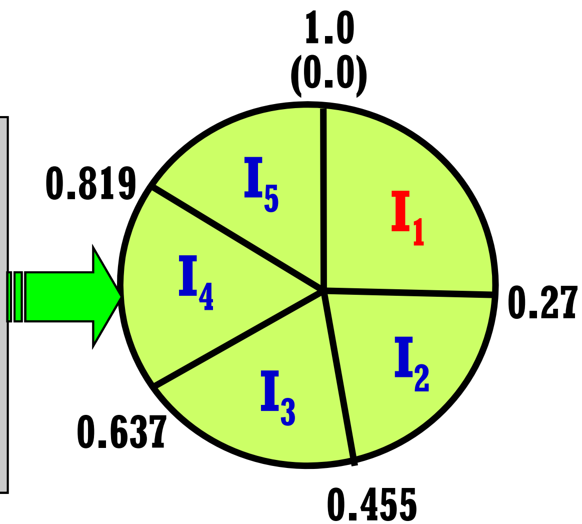
Scaling

- Relax the premature convergence
- Continual re-scaling is needed

$$f = a \cdot fitness + b$$

$$\overline{f} = \overline{fitness} \qquad f_{\max} = \phi \cdot \overline{f}$$

f
27
18.5
18.2
18.2
18.1



f Roulette Wheel

Ranking Selection

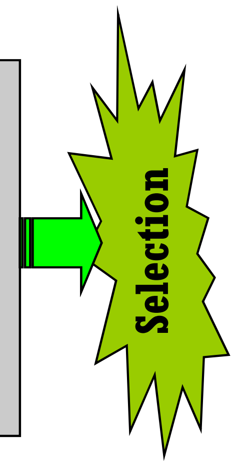
- Ranks of individuals are used
- Constant selection pressure

$$P[k] = \alpha + \beta \cdot k$$

With random number r

$$k = \frac{-(2\alpha+\beta)+\sqrt{(2\alpha+\beta)^2+8\beta \cdot r}}{2\beta}$$

	Fitness	Rank
I_1	80	5
I_2	10	4
I_3	5	3
I_4	3	2
I_5	2	1



Ordinal Selection (1)

Ranking Selection

- **Why** do we **resort to fitness** values?
  - The premature convergence has been brought forth from the fitness value itself
  - A key point in the selection is the **relative dominance** (i.e., **ranking**)!
- Ranking may lose some information, but simpler and more efficient

	ith Population	Fitness	Rank
I_1	101 … 110	80	5
I_2	111 … 100	10	4
I_3	001 … 100	5	3
I_4	010 … 111	3	2
I_5	101 … 011	2	1

- Suppose that the prob. of selecting the kth-rank individual

$$P[k] = \alpha + \beta \cdot k$$

- To be a probability distribution

$$\sum_{k=1}^{N} \alpha + \beta \cdot k = N\left(\alpha + \beta \frac{N+1}{2}\right) = 1$$

How to select individuals?
What criterion can be used for selection?

Ordinal Selection (2)

Ranking Selection (Cont.)

- Selection pressure is defined by

$$\phi = \frac{P[\text{selecting the fittest individual}]}{P[\text{selecting average individual}]}$$

$$\Rightarrow \quad \phi = \frac{\alpha + \beta \cdot N}{\alpha + \beta(N+1)/2}$$

$$\Rightarrow \quad \beta = \frac{2(\phi - 1)}{N(N-1)}, \quad \alpha = \frac{2N - \phi(N+1)}{N(N-1)}$$

which implies that $1 \le \phi \le 2N/(N+1) \simeq 2$

- The cumulative prob. distribution can be stated in terms of the sum of an arithmetic progress.

- With a random number r, finding the k is given by

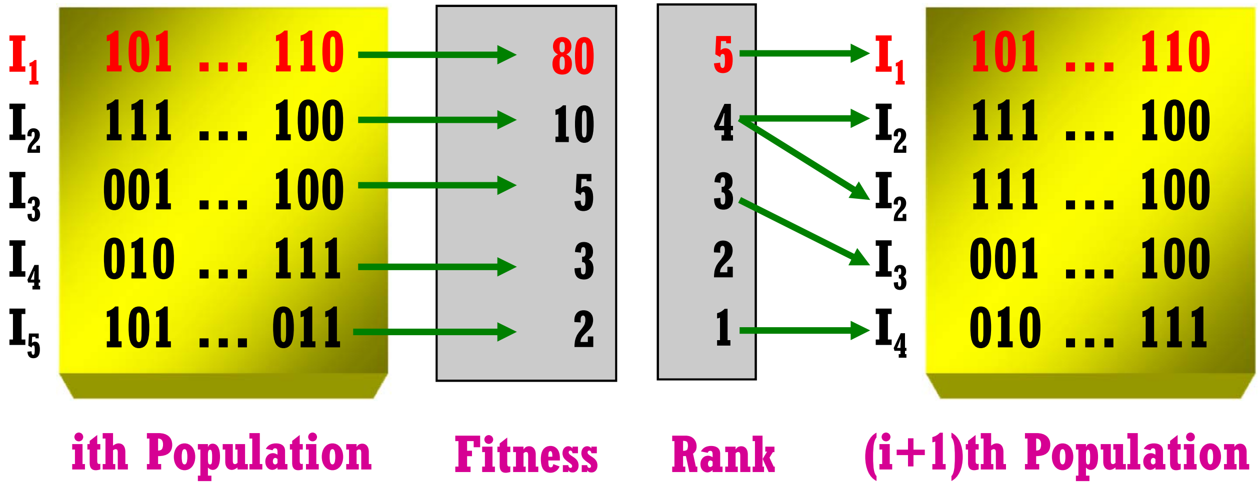
$$\sum_{i=1}^{k} (\alpha + \beta \cdot i) = \alpha \cdot k + \beta \frac{k(k+1)}{2} = r$$

$$\Rightarrow \quad k = \frac{-(2\alpha + \beta) + \sqrt{(2\alpha + \beta)^2 + 8\beta \cdot r}}{2\beta}$$

➔ With a given random number, the individual of the $\lceil k \rceil$-th rank can be selected!

Ordinal Selection (3)

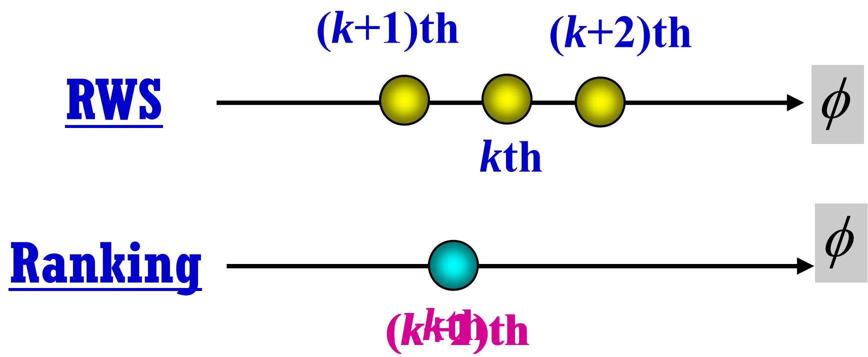
Ranking Selection (Cont.)



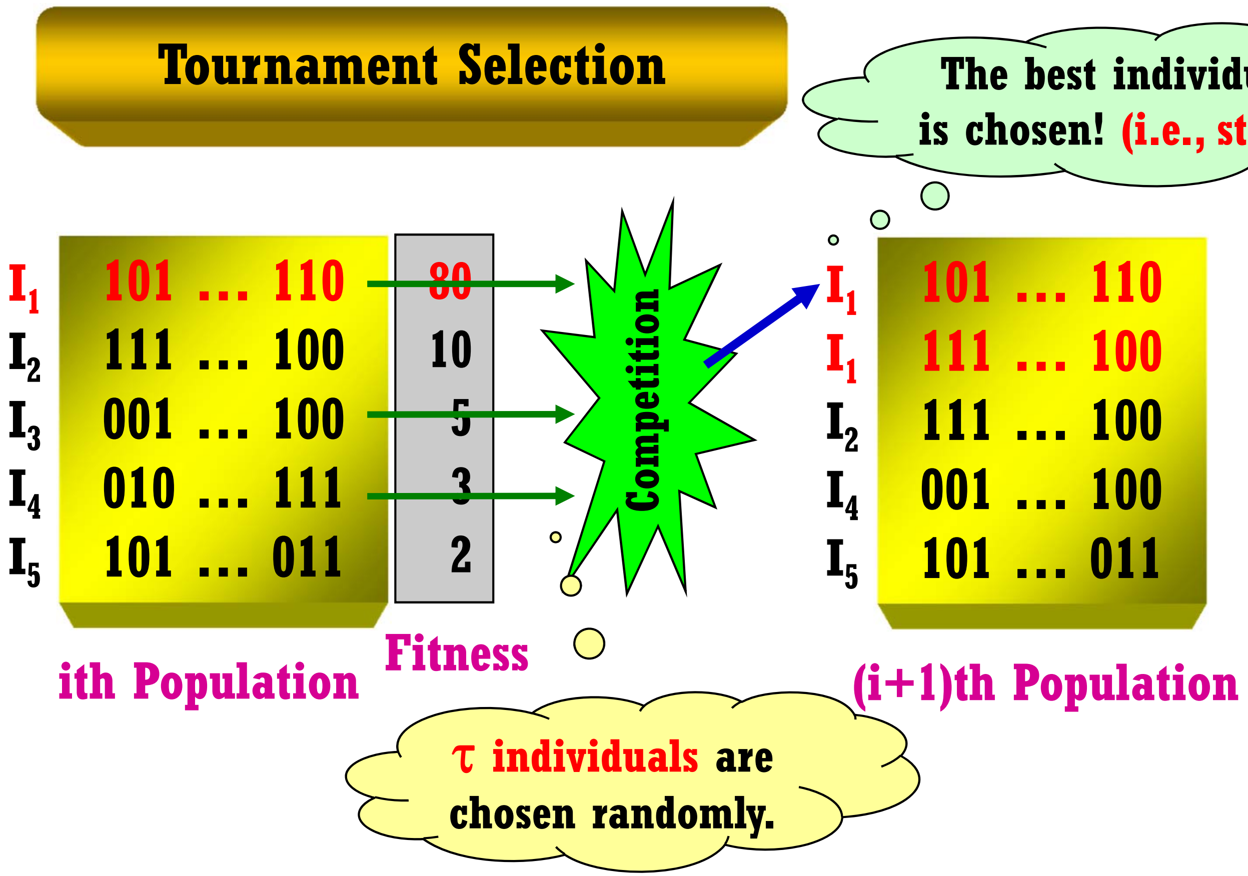
- Using the ordinary proportional selection, it takes *O(N)*
- But the ranking takes *O(NlogN)* by the sorting algorithm.
- Nevertheless, the prob. of keeping a constant selection pressure without re-scaling is an attractive one!

- For $N=5$, $\phi=1.5$, we can get $\alpha=\beta=1/20$

- For given random numbers (0.5, 0.9, 0.4, 0.1, 0.7), we have k=(3.22, 4.68, 2.77, 1.0, 4.0)
  ➔ (4, 5, 3, 1, 4) individuals are selected!



Ordinal Selection (4)



- In a complete cycle (i.e., generating N individuals), each individual will be compared τ times on average

- Every time it is compared, the best one is selected all the time.
(it is called 'strict' tournament.)

$$P[selecting\ I_{best} \mid I_{best}\ is\ chosen] = 1$$

$$\phi = \frac{P[selecting\ I_{best}]}{P[selecting\ I_{avg}]}$$

$$= \frac{P[selecting\ I_{best} \mid I_{best}\ is\ chosen]P[I_{best}\ is\ chosen]}{P[selecting\ I_{avg} \mid I_{avg}\ is\ chosen]P[I_{avg}\ is\ chosen]}$$

Ordinal Selection (5)

Tournament Selection (Cont.)

I_1 101 … 110
… … … …
I_{avg} 001 … 100
… … … …
I_N 101 … 011

ith Population

Probability of choosing any individual below the median is 1/2

- The chance of the median individual being surviving is the prob. that the remaining (τ-1) ones are all worse: $(½)^{\tau-1}$

$$P[\text{selecting } I_{avg} \mid I_{avg} \text{ is chosen}] = (1/2)^{\tau-1}$$

- Since the probability of selecting each individual is equally likely as 1/N, we get the following result regardless of τ

$$P[I_{best} \text{ is chosen}] = P[I_{avg} \text{ is chosen}]$$

- Thus, the selection pressure is $\phi = 2^{\tau-1}$
- To obtain a selection pressures below 2, soft tournament can be used such that the chance of winning of the best is $p<1$
- We can get $\phi = 2^{\tau-1}\, p$
- The pair-wise soft tournament selection can produce the selection pressure as in the ranking: $\phi = 2\, p$ that exists [0, 2]

Comparison of Selection Pressures

Proportional Selection

$$\phi = \frac{P[selecting\ I_{best}]}{P[selecting\ I_{avg}]} = \frac{fitness^{(best)}}{\frac{1}{N}\sum_{k=1}^{N} fitness^{(k)}}$$

- The selection pressure varies each generation!
- We never control the selection pressure.

Scaling

$$f_{\max} = \phi \cdot \overline{f} \quad \Rightarrow \quad \phi = f_{\max} / \overline{f}$$

$$\phi = f_{\max} \Big/ \frac{1}{N}\sum_{k=1}^{N}(a \cdot fitness^{(k)} + b)$$

- Also, the selection pressure varies each generation!
- But, it can somehow control the selection pressure.
- To do this, re-scaling is needed every generation.

Ranking Selection

$$\phi = \frac{\alpha + \beta \cdot N}{\alpha + \beta(N+1)/2}$$

$$\beta = 2(\phi - 1)/N(N-1), \quad \alpha = \frac{2N - \phi(N+1)}{N(N-1)}$$

- The selection pressure is not altered!
- It controls the pressure without any re-scaling.
- But, the selection pressure only exists in [1, 2]

Tournament Selection

I. Strict Tournament

$$\phi = 2^{\tau - 1}$$

II. Soft Tournament

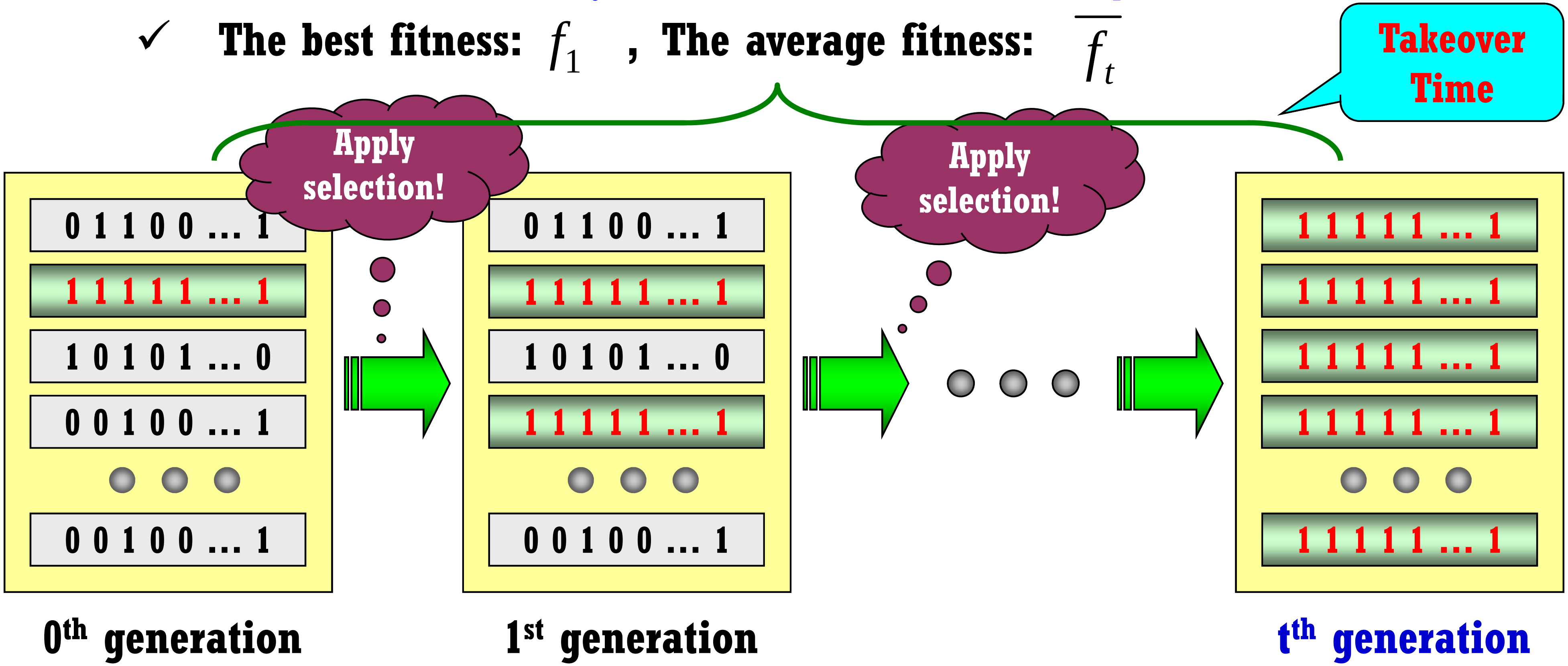
$$\phi = 2^{\tau - 1} p$$

- The selection pressure does not vary!
- It controls the pressure without any re-scaling.
- But, the selection pressure can have any number.

Takeover Time (1)

- **Takeover Time (Convergence Analysis by Applying Selection Alone)**
  - **Time from (an) initial best individuals until the population is converged**
  - **Assumption:**
    - ✓ Two alternatives: 0 and 1
    - ✓ Initial proportion $P_0=1/N$, Final proportional $P_f=(N-1)/N$
    - ✓ The best fitness: f_1 , The average fitness: $\overline{f_t}$

Worst-case Scenario: There is only one best individual!



Takeover Time (2)

Proportional Selection

You don't have to follow up all the derivations in detail.

- Suppose we have some number of distinct individuals with objective function values $f_j, j \in J$

- The proportion of the ith individual at time t ($P_{i,t}$) is related to the initial proportion of the other individuals and their function values as follows:

$$P_{i,t} = \frac{f_i^t P_{i,0}}{\sum_{j \in J} f_j^t P_{j,0}}$$

- We restrict ourselves to the unit interval and track the proportion of the best individuals, $P_{Best,t}$, where Best={x: $1 - 1/N \le x \le 1$}

$$P_{Best,t} = \frac{\int_{1-N^{-1}}^{1} f^t(x) p_0(x) dx}{\int_0^1 f^t(x) p_0(x) dx}$$

- With p_0=constant, and $f(x)=x^c$, we get

$$P_{Best,t^*} = 1 - (1 - 1/N)^{ct^*+1}$$

- Assuming a final proportion of best individuals $P_f = (N-1)/N$, we have

$$t^* = \frac{1}{c}(N \ln N) = \boxed{O(N \ln N)}$$

Takeover Time (3)

Tournament Selection — i.e., Ordinal Selection

- τ-wise tournament selection is considered:
  Draw τ individuals, and Select the best one!

I_1	101 … 110	P_t
I_1	101 … 110	
…	… … …	
I_N	101 … 011	

ith Population

➔ P_{t+1} equals the probability that the best individual is picked and survives (by τ-wise tournament) under P_t

- The case the best individual is copied into the next generation:
  If the best individual is drawn at least one time among τ times,

$$P_{t+1} = 1-(1-P_t)^\tau$$

$$1-P_{t+1} = (1-P_t)^\tau$$

- We use the complementary proportion $Q_t = 1 - P_t$

$$Q_{t+1} = Q_t^\tau \Longrightarrow Q_t = \underbrace{((((Q_0)^\tau)^\tau \cdots)^\tau)^\tau}_{t} = Q_0^{\tau^t}$$

- By the complementary proportion, $Q_0=(N-1)/N$, $Q_f=Q_{t^*}=1/N$ and taking the natural log:

$$\ln(1/N) = \tau^{t^*}\ln((N-1)/N)$$

- Recognizing that $\ln(1-x) \cong -x$ for small x

$$-\ln N = \tau^{t^*}\ln(1-1/N)$$

$$\ln N = \tau^{t^*}(1/N)$$

- Taking the natural log again:

$$\ln\ln N = t^*\ln\tau - \ln N$$

$$t^* = \frac{\ln N + \ln\ln N}{\ln\tau} = \boxed{O(\ln N)}$$

Summary

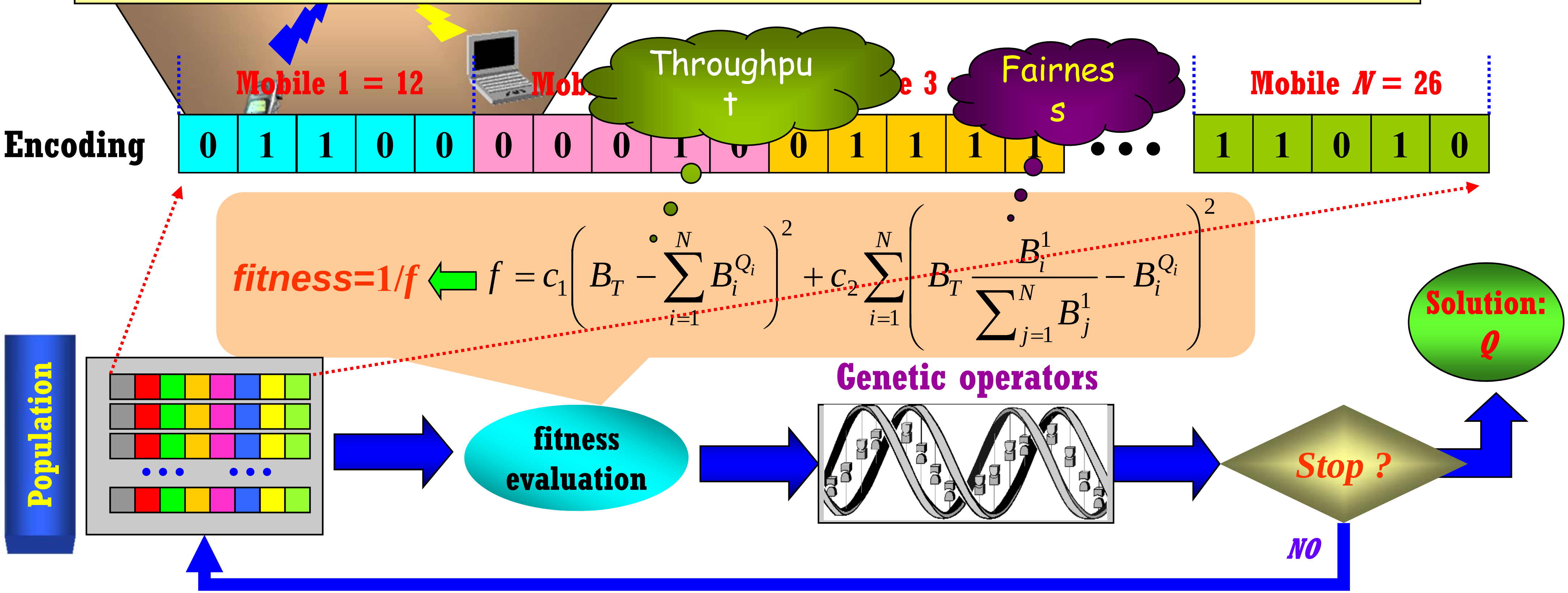
- There are **Two Selection Categories**!
  - **Proportional Selection**; e.g., Roulette-Wheel selection, Scaling
  - **Ordinal Selection**; e.g., Ranking selection, Tournament selection, etc.
- Generally, **Proportional Selection** tends to have **Premature Convergence**
  - Thus, the **scaling method** has been employed.
  - But, it **needs** to do **re-scaling** at every generation.
- **Ordinal Selection** is quite **robust** in this regard.
  - It can adjust **selection pressure** at a **constant level** what we want.
  - But, the **ranking selection** is somewhat **restricted**.
  - **Tournament selection does not have** such constraints.
- In terms of **takeover time** (i.e., convergence with selection only)
  - **Proportional Selection** has **O(NlnN)**, but **Ordinal Selection** has **O(lnN)**.

Thus, we conclude that the **Tournament Selection is the most promising choice**!

Some Real-World Applications

Application: Resource Allocation

QoS Index	1	2	…	14	15	…	26	27
Video	High	High	…	Mid.	Mid.	…	Low	Low
Audio	High	High	…	Mid.	Mid.	…	Low	Low
Data	High	Mid.	…	Mid.	Low	…	Mid.	Low



$$\text{fitness}=1/f \Leftarrow f = c_1\left(B_T - \sum_{i=1}^{N} B_i^{Q_i}\right)^2 + c_2 \sum_{i=1}^{N}\left(B_T \frac{B_i^1}{\sum_{j=1}^{N} B_j^1} - B_i^{Q_i}\right)^2$$

Time-Series Forecasting

What Is It?

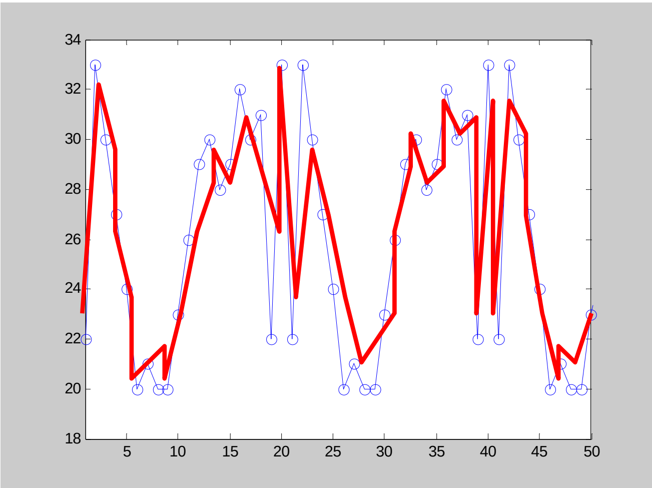
- Predicting some future outcomes from a set of historical events
- Stock prediction, Weather forecasting, Passenger prediction, etc.

Many Classical Methods Exist!

- Linear model approach: exp. smoothin ... MA
- Non-linear model approach: Threshold, GR...

Not so efficient!

Actually, it is not like this in real-coded domain!



GAs: linear model

GA Approach:

- Using a linear-type function: i.e., Future can be represented by a linear combination of past data.

$$x_{t+1} = \alpha_t x_t + \alpha_{t-1} x_{t-1} + \alpha_{t-2} x_{t-2} + \cdots + \alpha_{t-6} x_{t-6}$$
$$= \sum_{k=0}^{6} \alpha_{t-k} x_{t-k}$$

Encoding

α_t	α_{t-1}	α_{t-2}	α_{t-3}	α_{t-4}	α_{t-5}	α_{t-6}
0.83	0.57	0.25	0.92	0.23	0.41	0.55
0.51	0.13	0.46	0.19	0.88	0.76	0.83

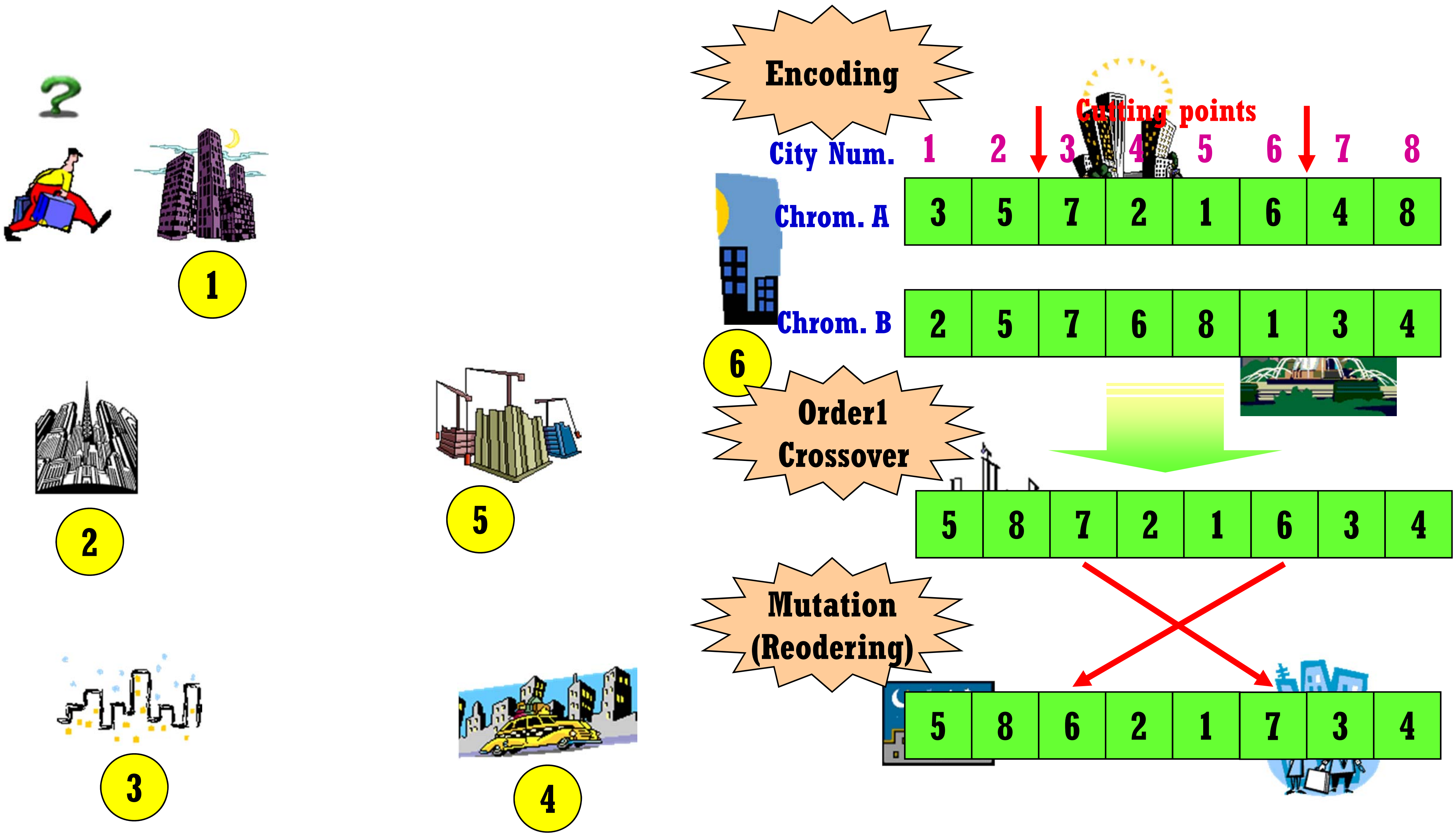
Crossover

0.83	0.57	0.46	0.19	0.88	0.76	0.55
0.51	0.13	0.25	0.92	0.23	0.41	0.83

Mutation

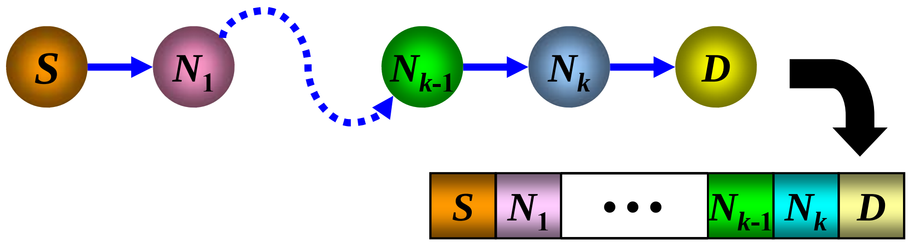
0.51	0.13	0.25	0.92	0.16	0.41	0.83

Traveling Salesman Problem

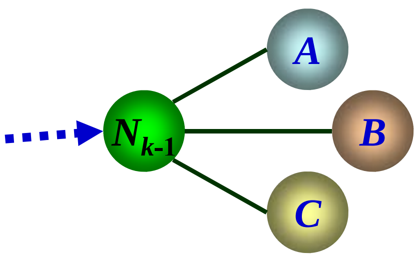


Shortest Path Routing

Encoding: X-ary representation



Initialization: Random

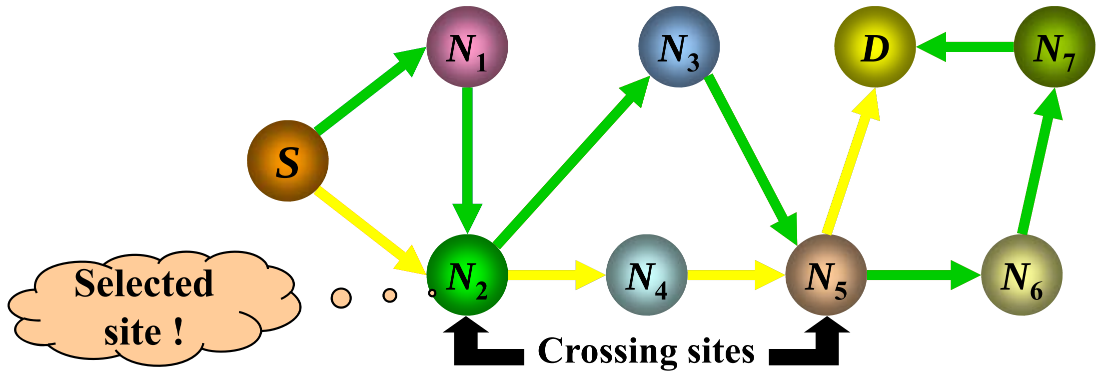


$$P[N_{k-1} \rightarrow A] = P[N_{k-1} \rightarrow B] = P[N_{k-1} \rightarrow C]$$

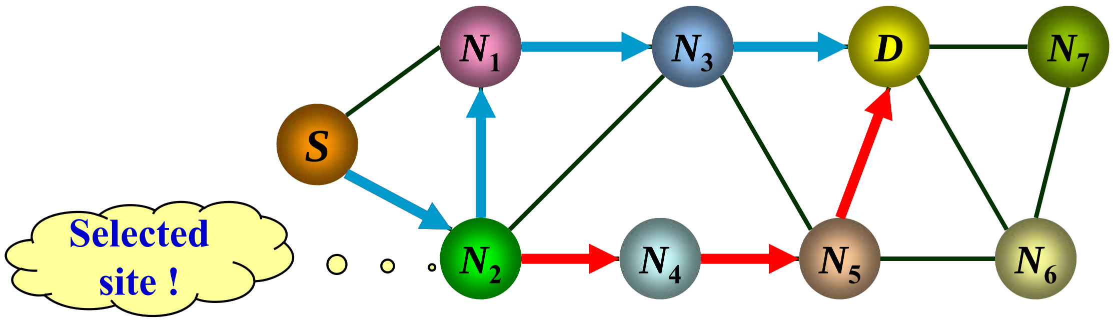
Fitness $F_i=[\sum C(g_i(j),g_i(j+1))]^{-1}$

Selection: Proportional & Ordinal

Crossover: Single to Uniform



Mutation: Perturbation



Treating Infeasible Solutions: Penalty & Repair