Studi Kasus

Dr. Suyanto, S.T., M.Sc. HP/WA: 0812 845 12345

Intelligence Computing Multimedia (ICM)
Informatics faculty – Telkom University

Intro

- "Untuk masalah seperti ini, apakah saya bisa menggunakan EAs?
- Kalau ya, algoritma EAs yang mana yang harus saya gunakan?
- Operator evolusi mana yang paling tepat?
- Bagaimana melakukan setting parameter yang baik?"

Strategi menggunakan EAs

- Kenali masalah yang anda hadapi
- Ruang masalah: seberapa besar?
- Analisa matematika: ada atau tidak?
- Jenis solusi: harus paling baik?
- Batasan waktu: ada atau tidak?
- Biner, Real, Automata atau Program?
- Nilai optimum atau permutasi?
- Ada pengetahuan untuk masalah yang dihadapi?

Strategi menggunakan EAs

- Setting parameter EAs
 - Observasi parameter
 - Pengontrolan parameter secara adaptif

Studi Kasus

- Peramalan Data Time Series
- TSP dengan batasan
- Penjadwalan kuliah

Peramalan Data Time Series

- Misalkan kita ingin membangun sistem peramalan hasil penjualan komputer di suatu perusahaan vendor besar.
- Misalkan kita memiliki data-data penjualan harian dari tanggal 01 Januari 2006 sampai 14 Desember 2007 yang ditunjukkan pada tabel berikut.

Hasil penjualan komputer per hari (01 Jan 2006 - 14 Dec 2007)

Tanggal	Penjualan (miliar rupiah)
14 Dec 2007	99,9573
13 Dec 2007	99,8459
12 Dec 2007	98,8708
11 Dec 2007	98,7480
10 Dec 2007	98,3897
09 Dec 2007	97,6780
08 Dec 2007	97,3797
	•••
03 Jan 2006	90,7597
02 Jan 2006	90,5770
01 Jan 2006	89,3897

Peramalan untuk Data Time Series

- Pada kasus peramalan ini, kita bisa melihat bahwa peramalan yang dimaksud adalah peramalan menggunakan data time series.
- Artinya, untuk memprediksi penjualan pada hari *H*, kita hanya menggunakan hasil-hasil penjualan pada hari-hari sebelumnya, *H*-1, *H*-2, dan seterusnya.
- Dengan demikian, masalah ini dapat dimodelkan secara linier sebagai berikut

Peramalan untuk Data Time Series

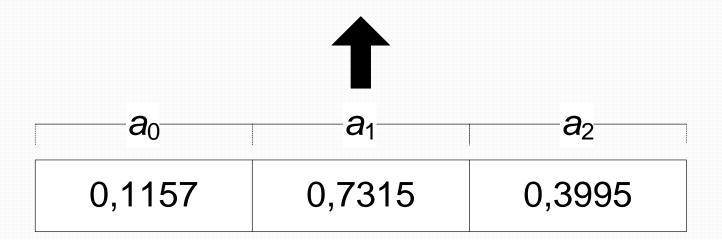
Dengan demikian, masalah ini dapat dimodelkan secara linier sebagai berikut

$$z = a_0 + a_1 y_1 + a_2 y_2 + \dots + a_k y_k$$

dimana y_1 sampai y_k adalah masukan yang berupa hasil-hasil penjualan pada hari-hari sebelumnya, H-1, H-2, ..., H-k.

Representasi individu GA

$$z = 0.1157 + 0.7315y_1 + 0.3995y_2$$



Bagaimana dengan GE?

```
N = \{ \text{expr, op, pre\_op} \}
T = \{Sin, Cos, Tan, Log, +, -, /, *, y1, y2, y3, y4, 0,5, 1, 1,5, 2, ()\}
S = \langle \exp r \rangle
P dapat direpresentasikan sebagai:
(1) <expr> ::= <expr> <op> <expr>
                                                       (A)
                                                       (B)
                   | (<expr> <op> <expr>)
                   <pre_op> (<expr>)
                                                       (C)
                    <var>
(2) <op> ::= +
                  (A)
                  (B)
                  (C)
                  (D)
(3) < pre_op > ::= Sin
                                     (A)
                                     (B)
                   Cos
                   Tan
                                     (C)
                   Log
                                     (D)
(4) < var > := y1
                                     (A)
                  | y2
                                     (B)
                                     (C)
                   y3
                                     (D)
                   y4
                                     (E)
                   0,5
                                     (F)
                                     (G)
                   1,5
```

(H)

2

Bagaimana dengan GE?

• Grammar tersebut dibuat dengan asumsi bahwa model-model peramalan hanya memperhatikan maksimum 4 data masukan, y_1 sampai y_4 . Variabel a_0 sampai a_4 diasumsikan hanya bisa bernilai 0,5; 1; 1,5 atau 2. Dengan grammar di atas, kromosom-kromosom pada GE bisa menghasilkan model-model peramalan yang sangat bervariasi, seperti:

$$z = 0.5 + 2y_1 - y_2 + (0.5 * y_3)$$

atau

$$z = 1.5 + \sin(y_1) - \log(y_2)$$

atau

$$z = \frac{0.5 + y_1}{(y_2 + (0.5 * y_3) - \cos(y_4))}$$

Fungsi Fitness

$$f = \frac{1}{(K+b)}$$

dimana *b* merupakan suatu bilangan yang dianggap sangat kecil untuk menghindari pembagian dengan o, sedangkan *K* adalah rata-rata kesalahan peramalan untuk semua data penjualan.

Fungsi Fitness

Kesalahan peramalan merupakan harga mutlak dari selisih hasil peramalan menggunakan model tersebut (z) dengan data penjualan yang sebenarnya (z^*) . Dengan demikian, rata-rata kesalahan peramalan dapat dituliskan sebagai

$$K = \frac{1}{N} \sum_{i=1}^{N} \left| z_i - z_i^* \right|$$

dimana *N* adalah jumlah semua data peramalan.

Traveling Salesman Problem

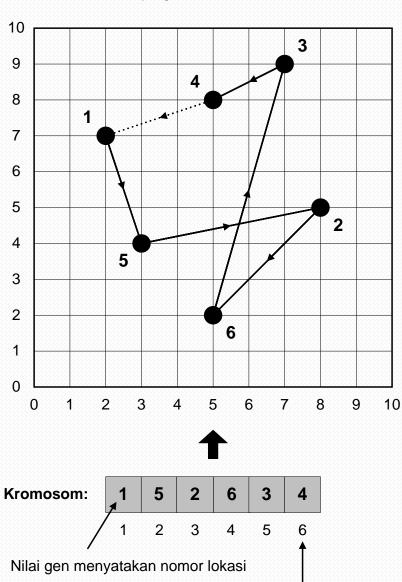
- Bagaimana menemukan urutan kunjungan
- Solusi → permutasi dengan biaya terendah
- Termasuk masalah Diskrit
- Algoritma EAs yang bisa dipakai?
- GA?
- ES, EP & DE ??
- GP & GE ???

TSP dengan GA

Fungsi fitness?

1 / total biaya

Individu: urutan kunjungan semua lokasi

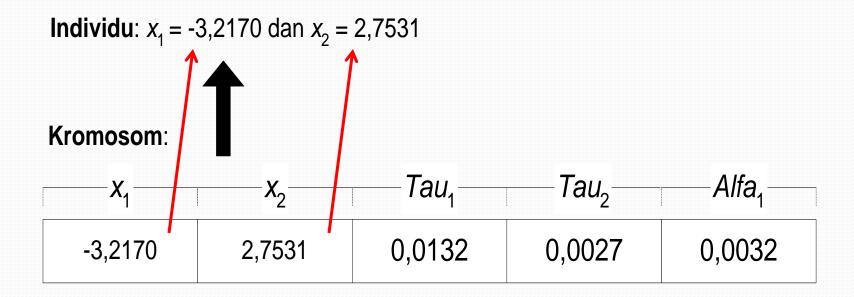


Posisi gen menyatakan urutan kunjungan

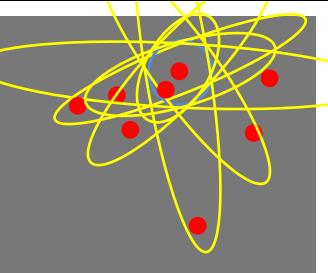
TSP dengan ES, EP dan DE

- Konversi kontinyu menjadi diskrit
- Variabel strategies
- Komputasi vektor pada DE

Kromosom

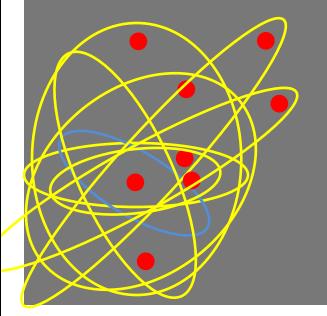


Mutasi dengan Korelasi, Pop=2, S = 7



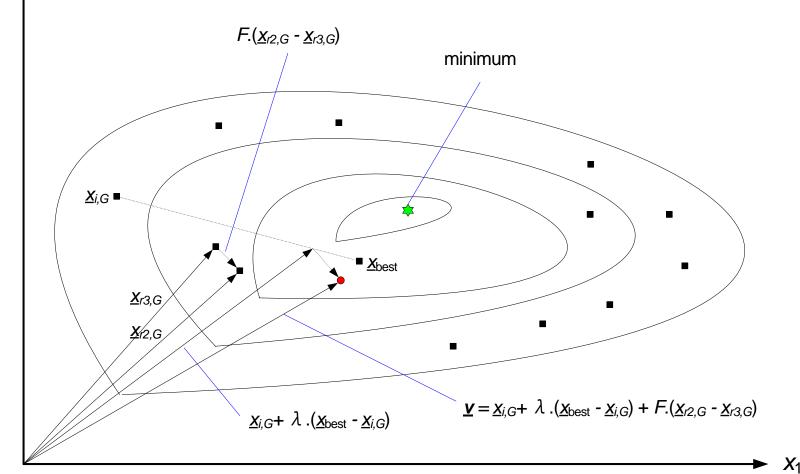


Global maximum

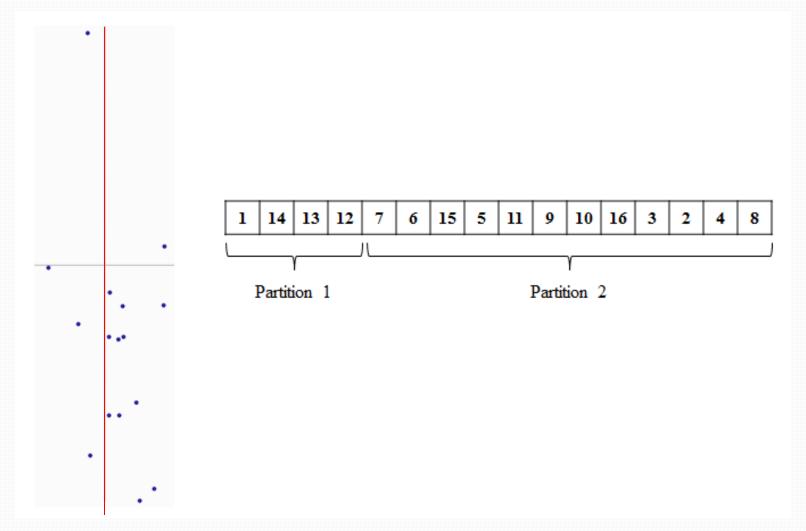


2 parent dan 14 anak hasil mutasi

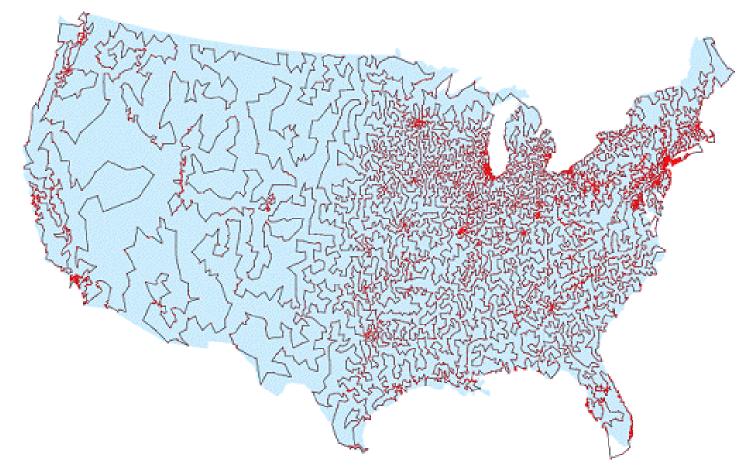
- Sejumlah NP vektor parameter pada generasi G
- Vektor parameter <u>v</u> baru yang dihasilkan
- Nilai minimum global yang dicari



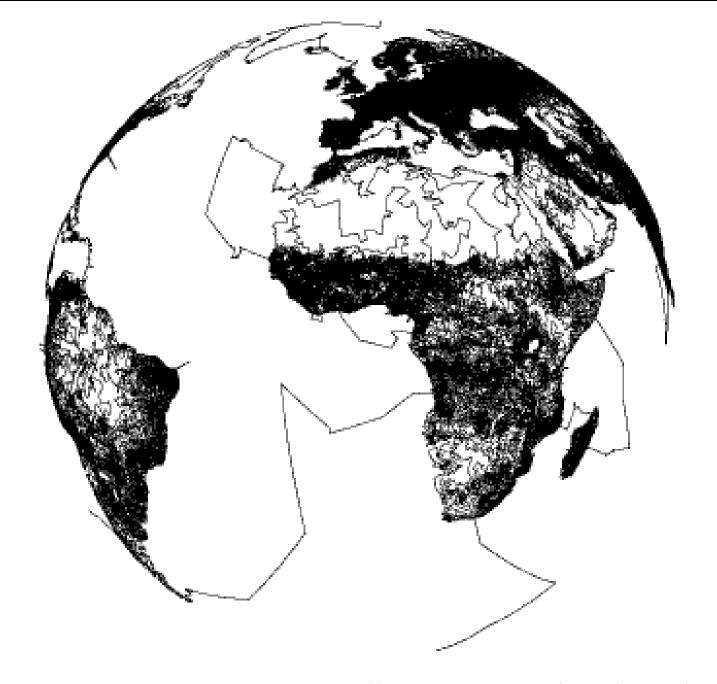
Two sub partitions by Y axis



13,509 cities in the US



http://www.nsf.gov/pubs/2002/nsf0120/nsf0120_usa.html



1,904,711-city instance of locations, http://www.tsp.gatech.edu/world/images/world.anim5a.gif

TSP dengan batasan

- Traveling Salesperson Problem (TSP) adalah suatu masalah optimasi kombinatorial.
- Pada kasus tertentu, mungkin saja terdapat batasan untuk kombinasi. Misalkan, kita ingin menemukan rute penerbangan dengan total jarak paling pendek untuk 60 kota besar di tiga benua berbeda, yaitu 20 kota di *North America*, 20 kota di *Europe*, dan 20 kota di *Sout-East Asia* (Jepang, Cina, Korea Selatan, dan Taiwan) [SUY05b].
- Untuk masalah ini, misalkan rute yang valid adalah rute yang pada setiap benua hanya mengunjungi maksimum dua kota (tidak boleh lebih).

Daftar 60 kota di tiga benua berbeda dengan posisi *Latitude* dan *Longitude*nya

No.	Nama kota	Latitude	Longitude	
1	Amarillo	35.207	- 101.834	
2	Atlanta 33.752		- 84.393	
3	Berkeley	37.869	-122.271	
4	Boston	42.356	-71.056	
5	Buffalo	42.881	-78.872	
	•••		•••	
60	60 Tianjin		117.20	

Inisialisasi populasi

- Untuk membangkitkan individu yang valid, kita bisa memilih secara acak satu atau dua kota dari 60 kota yang ada menggunakan dua aturan berikut ini:
 - Setelah meletakkan satu atau dua kota dari suatu benua, pilih secara acak satu atau dua kota dari benua yang lain.
 - Jumlah kota (satu atau dua kota) dari suatu benua dipilih berdasarkan suatu probabilitas, misalkan P_k, yang nilainya dalam interval [0, 1]. Jumlah kota yang dipilih adalah satu jika bilangan acak yang dibangkitkan adalah lebih kecil dari P_k. Tetapi, jika bilangan acak yang dibangkitkan adalah lebih besar atau sama dengan P_k, maka jumlah kota yang dipilih adalah dua.

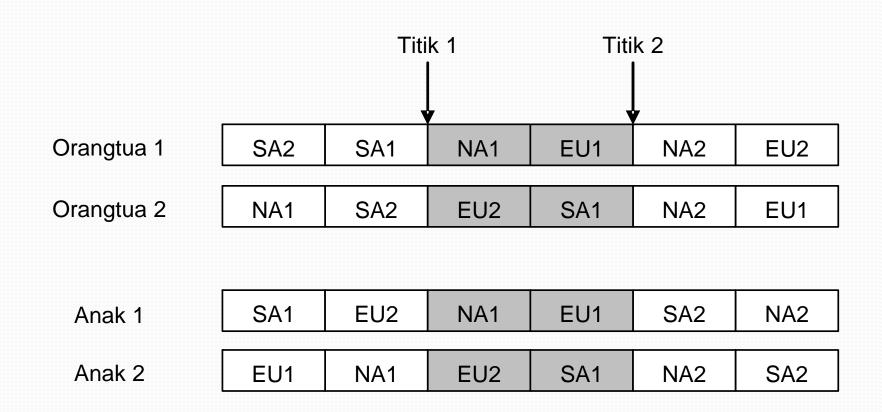
Inisialisasi populasi

	Gen 1	Gen 2	Gen 3	Gen 4	Gen 5	Gen 6		Gen 59	Gen 60
а	NA1	EU5	SA3	NA3	EU1	SA2	•••	EU11	SA9
b	NA1	NA6	EU1	EU5	SA5	SA1	•••	SA7	SA12
С	NA3	NA5	SA1	EU3	NA2	NA7	•••	EU4	EU7

Rekombinasi

- Untuk masalah TSP dengan batasan ini, kita bisa menggunakan rekombinasi untuk representasi permutasi, seperti: order crossover, partially mapped crossover, cycle crossover, maupun edge recombination.
- Tetapi, diperlukan sedikit modifikasi agar anak-anak yang dihasilkan tetap valid. Misalkan kita menggunkan order crossover. Modifikasi seperti apa yang harus kita lakukan?
- Kita bisa memberikan batasan bahwa peletakkan kota harus mengikuti aturan seperti pada inisialisasi.

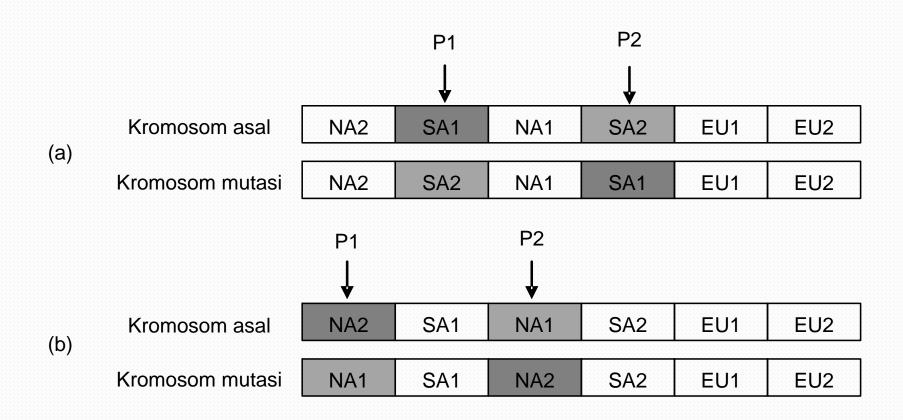
Rekombinasi



Mutasi

- Untuk operator mutasi, kita bisa menggunakan swap mutation yang diberikan sedikit batasan, yaitu jika dua gen yang dipertukarkan harus berupa kota yang berasal dari benua yang sama.
- Artinya, jika gen pertama yang terpilih adalah kota dari benua SA, maka gen kedua yang akan dipertukarkan harus berupa kota yang berasal dari SA juga.

Mutasi



Fungsi fitness

 Jarak tempuh antara dua kota dihitung berdasarkan perhitungan jarak yang disebut the great circle distance (untuk lebih detail, silahkan lihat di situs www.zipcodeworld.com):

$$d = 69.1 \left(\frac{180}{\Pi}\right) \arccos[\sin(Lat_1)\sin(Lat_2) + ...$$
$$\cos(Lat_1)\cos(Lat_2)\cos(Long_2 - Long_1)]$$

dimana *d* adalah jarak antara dua kota dalam *km*.

Fungsi fitness

Karena masalahnya adalah minimasi, maka fungsi *fitness*-nya adalah:

$$f = \frac{1}{c(d)}$$

Mengapa tidak menggunakan bilangan kecil *a*? Karena total biaya perjalanan tidak mungkin o.

Penjadwalan kuliah

- Penjadwalan kuliah merupakan suatu masalah yang sangat kompleks.
- Misalnya, suatu fakultas di sebuah universitas memiliki 36 ruang kuliah yang terdiri dari 24 ruangan besar dengan kapasitas di atas 50 mahasiswa dan 12 ruangan kecil dengan kapasitas 30 mahasiswa.
- Misalkan, setiap semester terdapat 400 pertemuan kuliah (setiap pertemuan membutukan waktu 2 jam) yang harus ditentukan jadwal kuliah dan ruangannya dalam satu minggu (Senin sampai Sabtu).

Penjadwalan kuliah

- Hari Senin sampai Kamis dan hari Sabtu, jam kuliah adalah 07.00-15.00 (8 jam). Hari Jum'at, jam kuliah adalah 07.00-11.00 dan 13.00-15.00 (6 jam). Sehingga jumlah jam kuliah dalam satu minggu adalah 46 jam. Karena setiap pertemuan kuliah membutuhkan waktu 2 jam, maka dalam satu minggu terdapat maksimum 23 pertemuan.
- Status dosen dibedakan menjadi dua, yakni dosen Dalam dan dosen Luar Biasa (LB). Seorang dosen mungkin mengajar lebih dari satu kelas kuliah.

Penjadwalan kuliah

Batasan-batasan yang digunakan:

- Tidak boleh terjadi bentrok, baik waktu maupun ruangan, untuk dosen maupun kelas kuliah yang sama.
- Dosen Dalam maupun dosen LB yang mempunyai jabatan akademik boleh meminta jadwal kuliah, dengan aturan bahwa dosen dengan jabatan akademik Guru Besar dan Lektor Kepala lebih diprioritaskan tiga kali lebih besar dibanding dosen dengan jabatan akademik Lektor dan Asisten Ahli. Jadwal yang dihasilkan sebisa mungkin dapat memenuhi permintaan tersebut.

Penjadwalan kuliah

Batasan-batasan yang digunakan:

- Jadwal kuliah untuk dosen Dalam, diutamakan hari Senin sampai Kamis dengan nilai prioritas empat kali lebih besar dibanding hari Jum'at dan Sabtu. Sedangkan untuk dosen LB, diutamakan hari Jum'at dan Sabtu dengan nilai prioritas empat kali lebih besar dibanding hari Senin sampai Kamis. Kecuali atas permintaan dosen yang bersangkutan (lihat aturan b di atas).
- Ruang kuliah yang besar diutamakan untuk kelas kuliah yang pesertanya lebih dari 30 mahasiswa. Nilai prioritasnya adalah empat kali lebih besar dibanding kelas kuliah yang pesertanya kurang dari 30 mahasiswa.

Penjadwalan kuliah

- Bagaimana menyelesaikan tersebut?
- Pertama kita lihat dulu seberapa besar ruang masalahnya? Dari 36 ruangan, 400 kelas kuliah, dan 23 pertemuan kuliah per minggu, maka jumlah solusi yang mungkin adalah sebanyak 400^(23 x 36).
- Suatu jumlah yang sangat besar jika harus diselesaikan dengan algoritma konvensional.
- Misalkan kita ingin menyelesaikan masalah ini menggunakan EAs. Dua hal yang harus kita definisikan adalah: representasi individu dan fungsi fitness-nya.

Representasi individu

Basis data Pertemuan Kuliah

No	Kode MK	Kode Kelas	Kuota mhs	Kode Dosen	Waktu	Ruangan
1	CS3143	IF-33-01	40	SBK		
2	CS2315	IF-34-01	35	TWG		
400	CS4923	IF-32-03	17	DSS		

Representasi kromosom

Ruang 1

	SN	SL	RB	KM	JM	SB
07-09	50		7	112	187	
07-09	32	1	101		400	
11-13		3		9		
13-15				6		
15-17	43	21	29	332		15

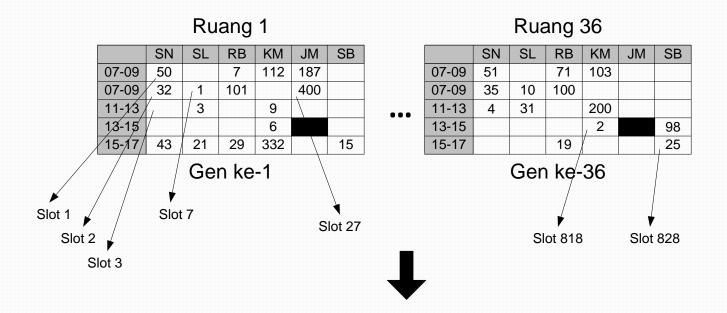
Gen ke-1

Ruang 36

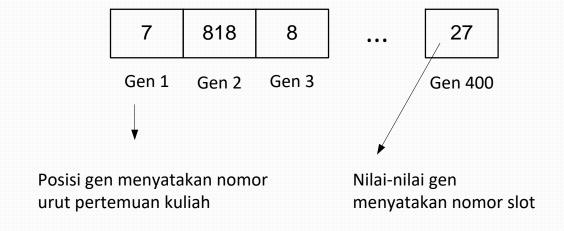
	SN	SL	RB	KM	JM	SB
07-09	51		71	103		
07-09	35	10	100			
11-13	4	31		200		
13-15				2		98
15-17			19			25

Gen ke-36

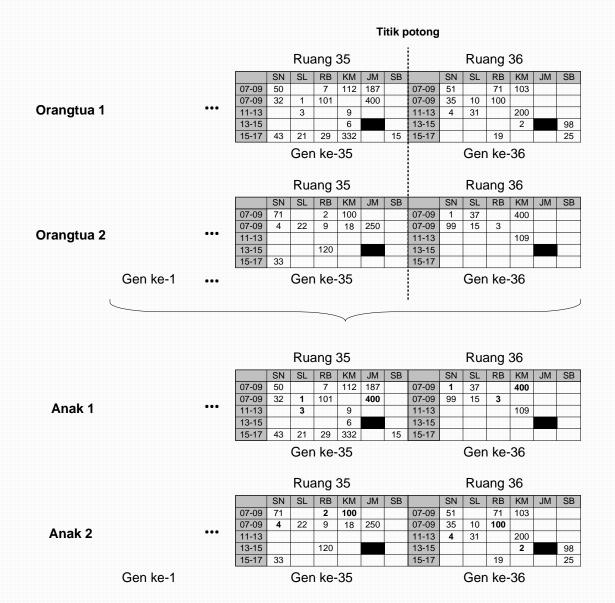
Kromosom dengan 36 gen:



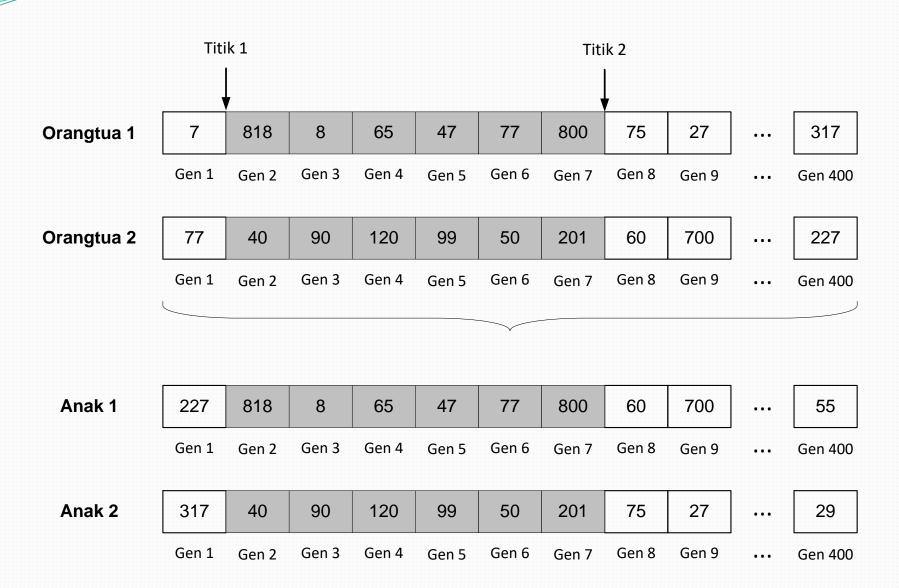
Kromosom dengan 400 gen:



Rekombinasi 1



Rekombinasi 2



Mutasi 1

Kromosom asal

Kromosom hasil mutasi

Ruang 1

	SN	SL	RB	KM	JM	SB
07-09	50		7	112	187	
07-09	/ 32	1	101		400	
11-13 /		3		9		
13-15				6		
15-17	43	21	29	332		15

Gen ke-1

Slot 1

Ruang 36

	SN	SL	RB	KM	JM	SB
07-09	51		71	103		
07-09	35	10	100			
11-13	4	31		200		
13-15				2		98
15-17			19			25

Gen ke-36



Ruang 1

	SN	SL	RB	KM	JM	SB
07-09			7	112	187	
07-09	32	1	101		200	
11-13		3		9		
13-15			50	6		
15-17	43	21	29	332		15

Gen ke-1

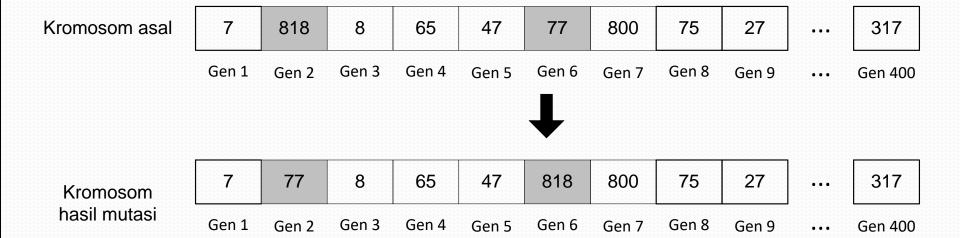
Slot 14

Ruang 36

	SN	SL	RB	KM	JM	SB
07-09	51		71	103		
07-09	35	10	100			
11-13	4	31		400		
13-15				2		98
15-17			19			25

Gen ke-36

Mutasi 2 (Swap)



Fungsi fitness

Batasan	Penalti	Alasan			
Tidak bentrok	24	Tidak bentrok merupakan suatu <i>hard constraint</i> (batasan keras) yang sebaiknya tidak dilanggar.			
Permintaan Guru Besar dan Lektor Kepala	12	Sehingga batasan ini lebih utama dibandingkan harus memenuhi permintaan dosen berjabatan akademik Guru Besar dan Lektor Kepala atau			
Permintaan Lektor dan Asisten Ahli	4	batasan-batasan yang lain. Misal batasan diberikan nilai prioritas sebesar 24. Kemudi			
Status Dosen	4	batasan permintaan dosen berjabatan akade Guru Besar dan Lektor Kepala diberi nilai prio			
Ruang Kuliah	4	sebesar 12. Misalkan ketiga batasan yang lainnya diberikan nilai priors yang sama, yaitu 4.			

Fungsi fitness

Pemberian Penalti atau pembobotan seperti pada tabel di atas belum tentu yang paling baik. Secara praktis, kita bisa melakukan observasi pembobotan tersebut selama melakukan *running* GA.

Fungsi fitness

$$f = \frac{1}{(B+a)}$$

$$B = J_a + J_b + J_c + J_d + J_e$$

- J_a = Jumlah jadwal pertemuan yang bentrok dikali Penalti (24),
- J_b = Jumlah jadwal yang tidak sesuai permintaan dosen berjabatan Guru Besar dan Lektor Kepala dikali Penalti (12),
- J_c = Jumlah jadwal yang tidak sesuai permintaan dosen Lektor dan Asisten Ahli dikali Penalti (4),
- J_d = Jumlah jadwal yang tidak sesuai ruang kuliahnya dikali Penalti (4)
- J_e = Jumlah jadwal yang Tidak Sesuai Status dosen Dalam atau Luar Biasa dikali Penalti (4).

Kesimpulan

- Sebelum memutuskan menggunakan EAs, sebaiknya suatu masalah dipahami secara benar.
- Keenam algoritma pada EAs memiliki karakteristik khusus yang harus diperhatikan dengan baik sehingga bisa memutuskan algoritma mana yang sebaiknya digunakan.

Daftar Pustaka

- [SUY08] Suyanto, 2008, Evolutionary Computation: Komputasi Berbasis "Evolusi" dan "Genetika", penerbit Informatika Bandung.
- [SUY05b] Suyanto and Sanaullah, 2005, "The Traveling Salesperson Problem Using Evolutionary Algorithm", technical report, Chalmers University of Technology, Sweden.