

# Evolving Fuzzy Systems

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

Intelligence Computing Multimedia (ICM)  
Informatics faculty – Telkom University

# Perbandingan EAs dan Fuzzy Systems

Kriteria	EAs	<i>Fuzzy Systems</i>
Sangat baik untuk masalah dengan informasi yang kurang presisi, tidak lengkap dan memiliki kebenaran parsial?	<b>Tidak</b>	Ya
Sangat baik untuk optimasi, khususnya permasalahan kombinatorial?	Ya	<b>Tidak</b>
Bisa <i>learning</i> ?	Ya	<b>Tidak</b>

# Permasalahan *Fuzzy Systems*

- Fungsi keanggotaan: **bentuk** & **batas2**
- Bagaimana membangun aturan *fuzzy*?

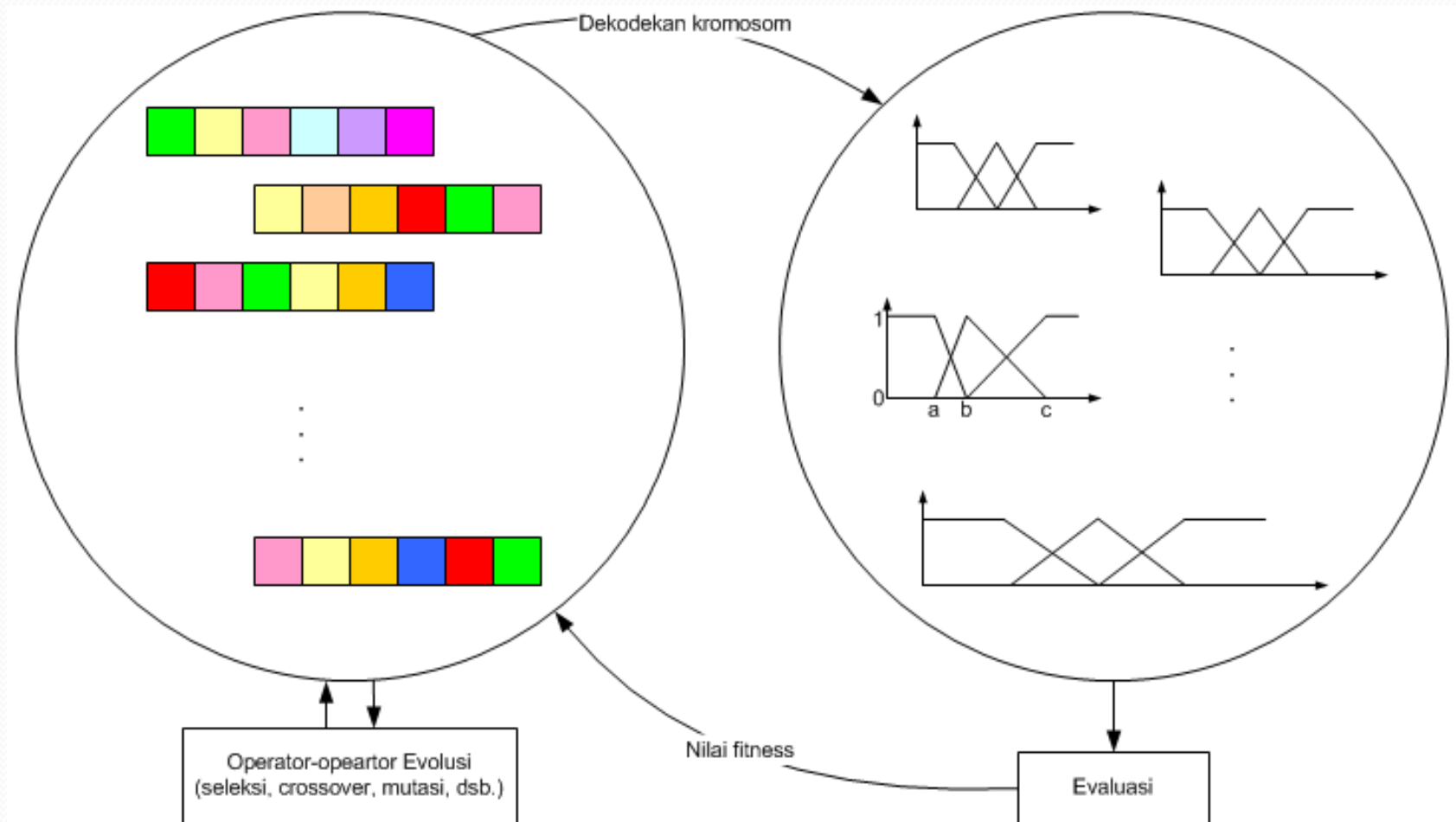
## Solusi?

Gunakan **EA** untuk mendapatkan fungsi keanggotaan dan/atau aturan *fuzzy* yang optimal.

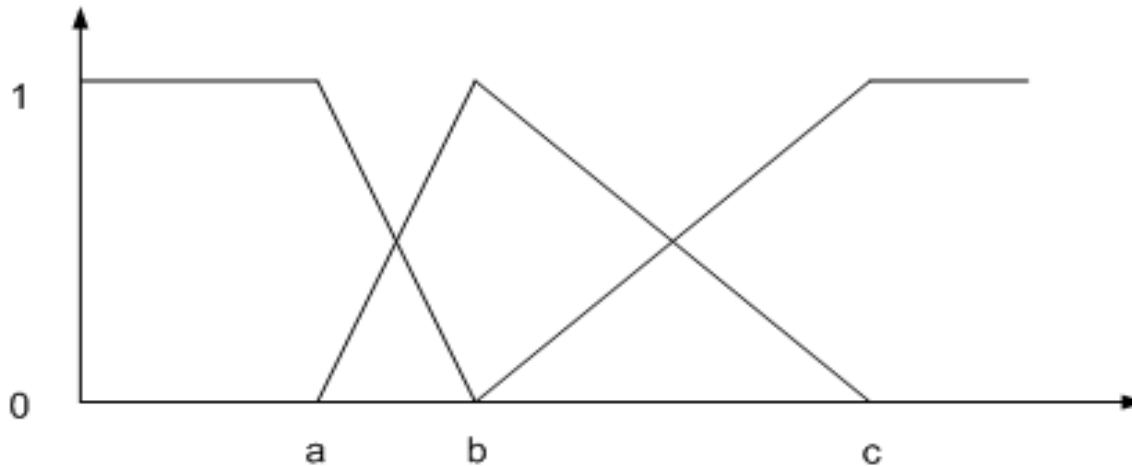
# *Tuning* fungsi keanggotaan

- Bentuk dan batas2 fungsi keanggotaan dibangkitkan berdasarkan data numerik.
- Untuk menghasilkan bentuk dan batas2 fungsi yang optimal, kita perlu melatih sistem dengan *training set* yang representatif.
- Untuk menghindari *overfit*, kita bisa menggunakan *validation set*.

# Tuning fungsi keanggotaan (statis)



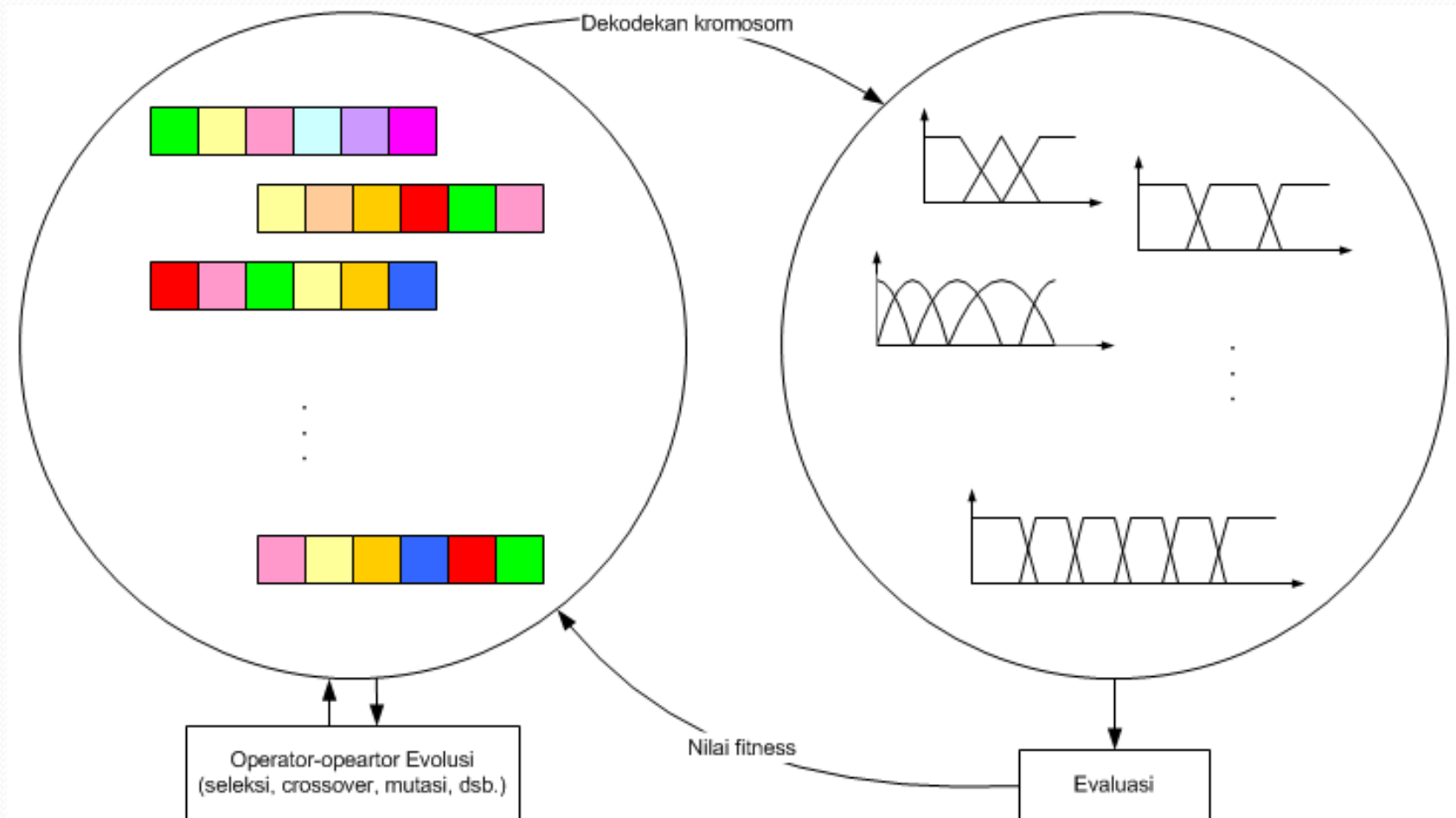
# Pengkodean Individu (statis)



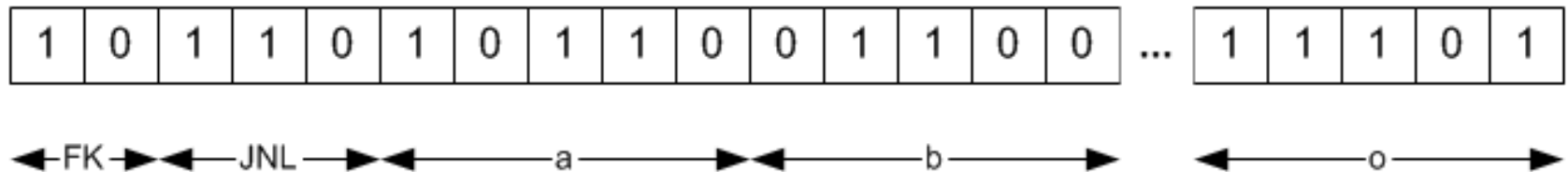
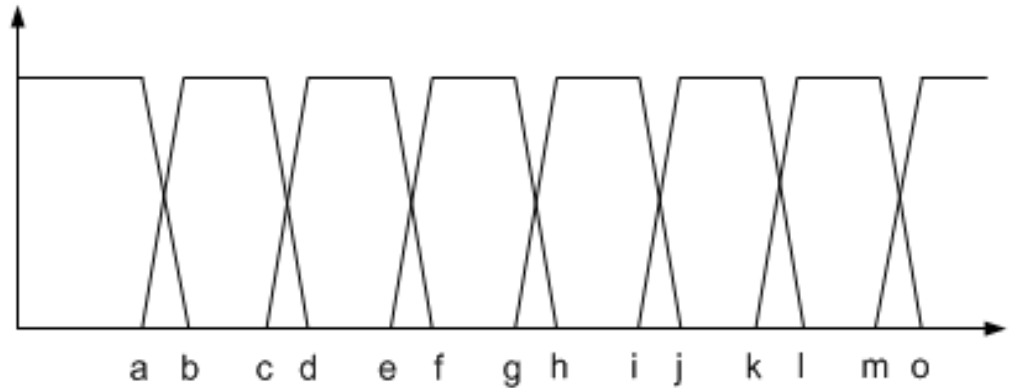
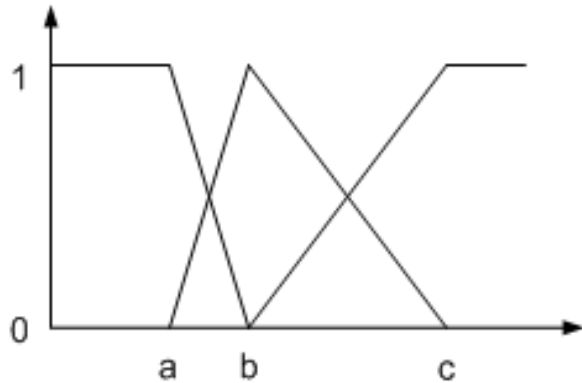
1	0	1	1	0	0	1	1	0	0	1	1	1	0	1
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

← a → → b → → c →

# *Tuning fungsi keanggotaan (dinamis)*



# Pengkodean Individu (dinamis)



FK: Bentuk Fungsi Keanggotaan (misal: segitiga, trapesium, phi, bell)

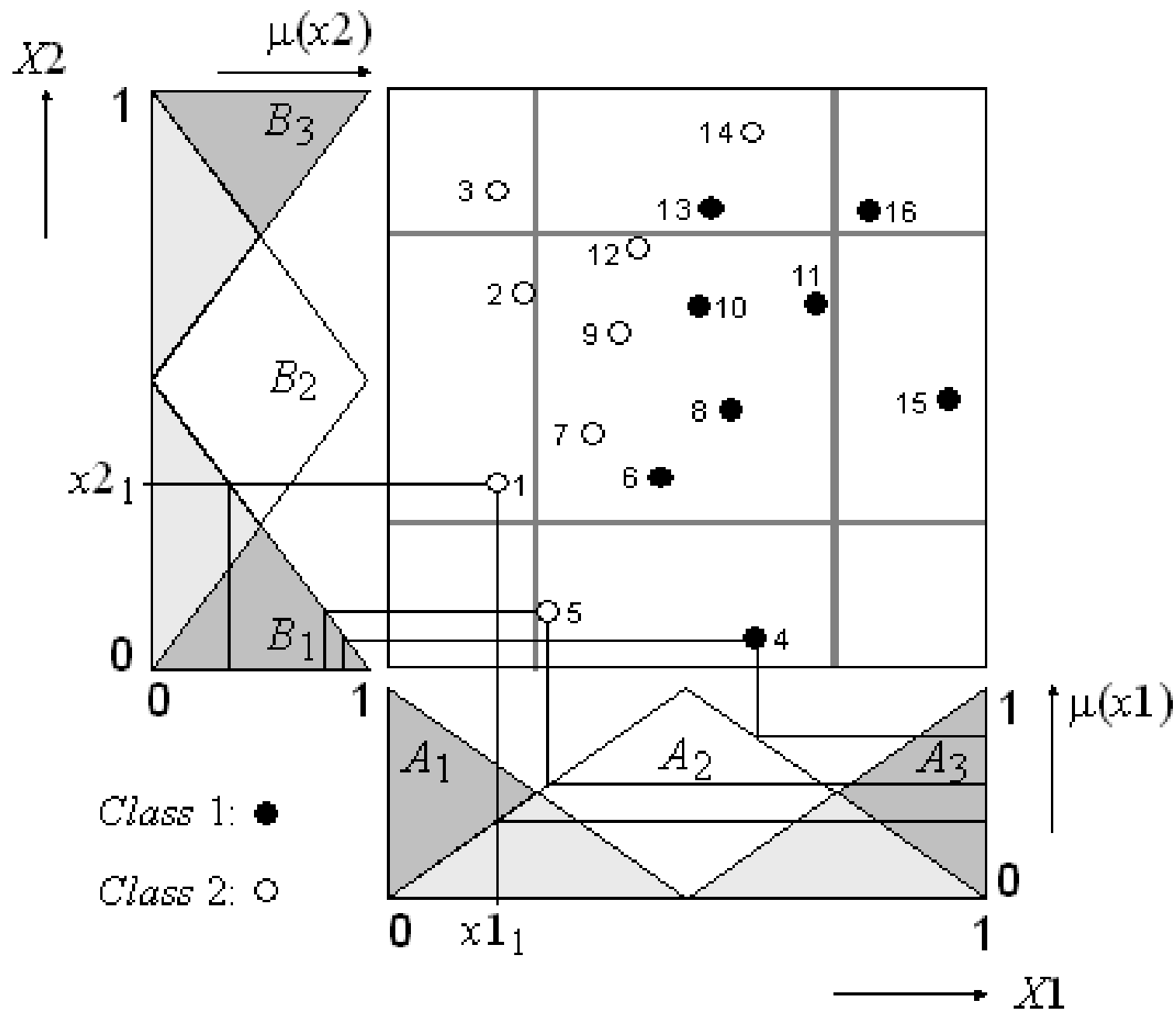
JNL: Jumlah Nilai Linguistik (misal maksimum  $2^3 = 8$ )

a, b, c, ..., o adalah batas2 fungsi keanggotaan



# Pemilihan Aturan *Fuzzy*

- Untuk mendapatkan himpunan aturan fuzzy, kita perlu melatih sistem dengan training set yang representatif dari sekumpulan data-data numerik.
- Untuk menghindari himpunan aturan fuzzy yang overfit, kita bisa menggunakan validation set.



# Fuzzy partition

- Titik-titik hitam dan putih menyatakan pola-pola dalam *Training Set* untuk *kelas 1* dan *kelas 2*.
- Kotak-kotak dapat dipandang sebagai **Tabel Aturan**.
- Nilai-nilai linguistik untuk input  $x_1$  ( $A_1$ ,  $A_2$  and  $A_3$ ) berada pada garis horizontal, dan nilai-nilai linguistik untuk input  $x_2$  ( $B_1$ ,  $B_2$  and  $B_3$ ) berada pada garis vertikal.
- Pada perpotongan baris dan kolom terdapat **rule consequent**.
- Pada Tabel Aturan, setiap sub ruang fuzzy hanya memiliki satu Aturan Fuzzy. Dengan demikian, jumlah aturan yang dapat dibangkitkan dari  $K \times K$  grid adalah sebanyak  $K \times K$ .

Aturan Fuzzy dalam  $K \times K$  fuzzy partition dapat direpresentasikan dalam bentuk [NEG02]:

*Rule  $R_{ij}$ :*

IF  $x1_p$  is  $A_i$   $i = 1, 2, \dots, K$

AND  $x2_p$  is  $B_j$   $j = 1, 2, \dots, K$

THEN  $\mathbf{x}_p \in C_n \left\{ CF_{A_i B_j}^{C_n} \right\}$   $\mathbf{x}_p = (x1_p, x2_p), p = 1, 2, \dots, P$

dimana  $x_p$  adalah pola latih pada ruang input  $X1 \times X2$ ,  $P$  adalah jumlah pola latih,  $C_n$  adalah rule consequent (kelas 1 atau 2), dan  $CF_{A_i B_j}$  adalah faktor kepercayaan bahwa suatu pola dalam sub ruang fuzzy  $A_i B_j$  termasuk ke dalam kelas  $C_n$ .

## Penentuan *rule consequent* [NEG02]

**Langkah 1:** Partisi ruang input ke dalam  $K \times K$  sub ruang fuzzy, dan hitung kekuatan masing-masing kelas pada pola-pola latih dalam setiap sub ruang fuzzy.

Kekuatan masing-masing kelas dalam sub ruang fuzzy direpresentasikan oleh jumlah pola-pola latihnya. Makin banyak pola latih, makin kuat kelas tersebut. Artinya, *rule consequent* untuk kelas tersebut menjadi lebih pasti.

**Langkah 2:** Tentukan *rule consequent* dan *certainty factor* pada setiap sub ruang fuzzy.

## Penentuan *rule consequent* [NEG02]

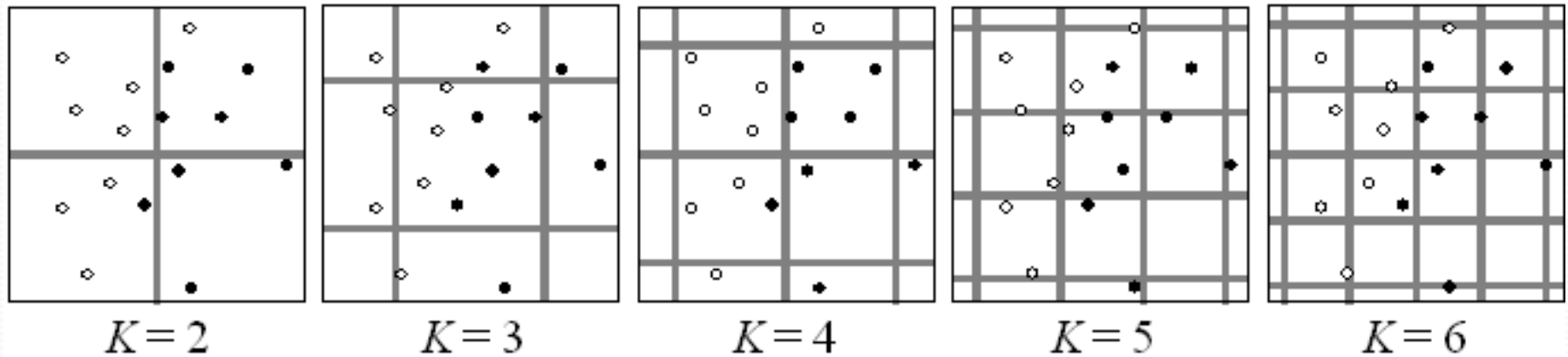
Certainty factor diinterpretasikan sbb:

- Jika semua pola latih dalam sub ruang fuzzy  $A_iB_j$  termasuk ke dalam kelas yang sama, maka *certainty factor*-nya adalah maksimum dan setiap pola baru dalam sub ruang tersebut pasti akan termasuk ke dalam kelas ini.
- Jika pola latih termasuk ke dalam kelas2 yang berbeda-beda dan kelas2 tersebut memiliki kekuatan yang sama, maka *certainty factor*-nya adalah minimum dan suatu pola baru akan termasuk ke dalam kelas mana saja. Hal ini berarti bahwa pola2 dalam sub ruang fuzzy dapat diklasifikasikan salah.
- Jika suatu sub ruang fuzzy tidak memiliki pola latih, maka kita tidak dapat menentukan *rule consequent*.
- Jika *fuzzy partition* terlalu kasar, maka banyak pola yang akan diklasifikasikan salah. Sebaliknya, jika *fuzzy partition* terlalu halus, banyak Aturan fuzzy tidak dapat dihasilkan karena kurangnya pola latih dalam masing2 sub ruang fuzzy.

## Langkah2 Penentuan *rule consequent* [NEG02]

- Pola-pola latih tidak selalu terdistribusi secara merata pada ruang input.
- Oleh karena itu, seringkali kita kesulitan untuk memilih ukuran *fuzzy grid* yang tepat.
- Untuk mengatasi kesulitan ini, kita bisa menggunakan Tabel Aturan Fuzzy Ganda atau ***multiple fuzzy rule tables***.

## Langkah2 Penentuan *rule consequent* [NEG02]



Aturan2 Fuzzy dibangkitkan untuk masing2 sub ruang fuzzy pada *multiple fuzzy rule tables*, sehingga kita mendapatkan suatu himpunan aturan yang lengkap sebanyak:

$$2 \times 2 + 3 \times 3 + 4 \times 4 + 5 \times 5 + 6 \times 6 = 90$$



## Langkah2 Penentuan *rule consequent* [NEG02]

Jika himpunan aturan  $S_{ALL}$  sudah dibangkitkan, suatu pola baru,  $\mathbf{x} = (x_1, x_2)$ , dapat diklasifikasi menggunakan prosedur sbb:

- **Langkah 1:** Pada masing-masing sub ruang fuzzy pada *multiple fuzzy rule tables*, hitung derajat kompatibilitas pola baru dengan masing2 kelas.
- **Langkah 2:** Tentukan dearajat kompatibilitas maksimum dari pola dengan masing2 kelas.
- **Langkah 3:** Tentukan kelas dimana pola baru memiliki dearajat kompatibilitas tertinggi, dan nyatakan bahwa pola baru adalah kelas tersebut.

## Langkah2 Penentuan *rule consequent* [NEG02]

- Jumlah *multiple fuzzy rule tables* yang diperlukan untuk suatu klasifikasi pola yang akurat mungkin sangat banyak
- Akibatnya, himpunan aturan yang dibangkitkan menjadi sangat banyak. Tetapi, aturan2 tersebut memiliki kemampuan klasifikasi yang berbeda2.
- Oleh karena itu, dengan memilih aturan2 yang potensial, kita dapat mengurangi jumlah aturan tetapi tingkat klasifikasi tetap tinggi.
- Dapatkah kita gunakan *Evolutionary Algorithms* (EAs) untuk menyeleksi Aturan Fuzzy tersebut?

## EAs untuk menyeleksi Aturan Fuzzy

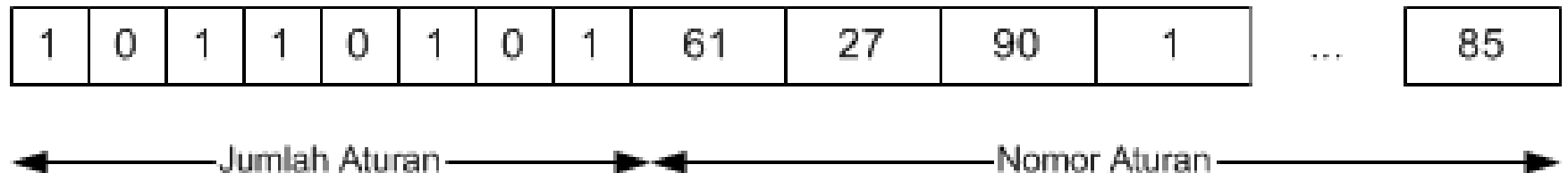
- Masalah pemilihan Aturan Fuzzy dapat dipandang sebagai masalah kombinatorial dengan **dua tujuan**.
- Tujuan pertama (paling penting) adalah memaksimalkan jumlah pola yang diklasifikasi dengan benar.
- Tujuan ke dua adalah meminimalkan jumlah aturan.

## EAs untuk menyeleksi Aturan Fuzzy

- Dua masalah paling mendasar dalam penggunaan EAs adalah:
  - Bagaimana bentuk individunya?
  - Bagaimana fungsi *fitness*-nya?
- Jika kedua masalah tersebut berhasil diselesaikan, maka komponen2 EAs yang lain kebanyakan sama dengan EAs untuk menyelesaikan masalah2 lainnya.

# Bentuk Individu

- Setiap gen menyatakan aturan2 dalam  $S_{ALL}$
- Jumlah aturan bersifat dinamis
- Terdapat gen yang menyatakan jumlah aturan



# Fungsi *Fitness*

Fungsi fitness harus mengakomodasi kedua tujuan tersebut, **memaksimalkan** jumlah pola yang diklasifikasi dengan benar dan **meminimalkan** jumlah aturan.

$$f(S) = w_P \frac{P_S}{P_{ALL}} - w_N \frac{N_S}{N_{ALL}}$$

dimana  $P_S$  adalah jumlah pola yang diklasifikasi dengan benar,  $P_{ALL}$  adalah jumlah pola latih.  $N_S$  dan  $N_{ALL}$  secara berturut-turut adalah jumlah Aturan Fuzzy dalam himpunan  $S$  dan  $S_{ALL}$ . Sedangkan  $w_P$  dan  $w_N$  adalah bobot untuk tujuan pertama dan tujuan ke dua.

# Fungsi *Fitness*

Jika memaksimalkan jumlah pola yang diklasifikasi dengan benar **lebih penting 5 kali** dibandingkan dengan meminimalkan jumlah aturan, maka fungsi *fitness*-nya menjadi:

$$f(S) = 5 \frac{P_s}{P_{ALL}} - \frac{N_s}{N_{ALL}}$$

# Studi Kasus

- *Sprinkler Control System*
- *Personalized Spam Filtering pada Discovery Challenge di ECML PKDD 2006, the 17th European Conference on Machine Learning and the 10th European Conference on Principles and Practice of Knowledge Discovery in Databases*



# *Sprinkler Control System*

Tanggal	Waktu	Suhu (°C)	Kelembaban Tanah (%)	Durasi Penyiraman (menit)
01-02-2006	08:00	20	16	55,7
01-02-2006	13:00	25	12	59,3
01-02-2006	18:00	16	30	5,6
02-02-2006	07:00	15	14	30,1
02-02-2006	12:00	27	10	43,4
02-02-2006	19:00	12	19	18,6
03-02-2006	06:30	19	12	22,1
03-02-2006	10:00	28	8	47,3
03-02-2006	13:30	35	6	76,4
03-02-2006	16:00	20	17	20,9

# *Sprinkler Control System*

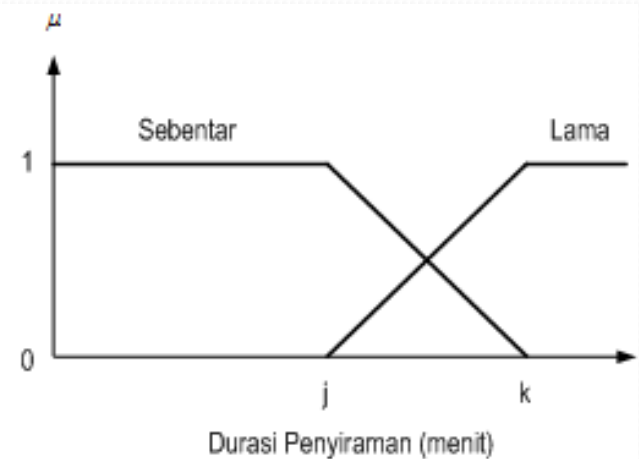
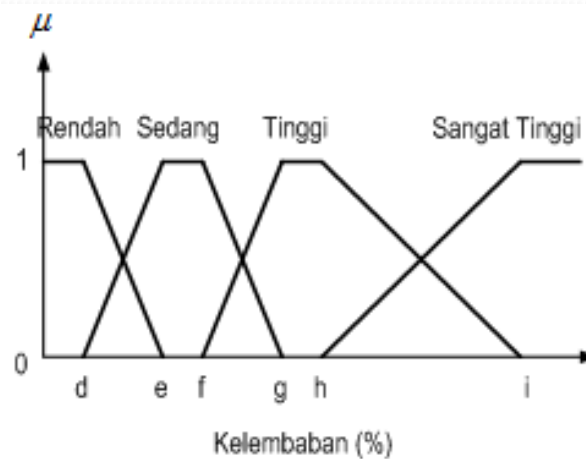
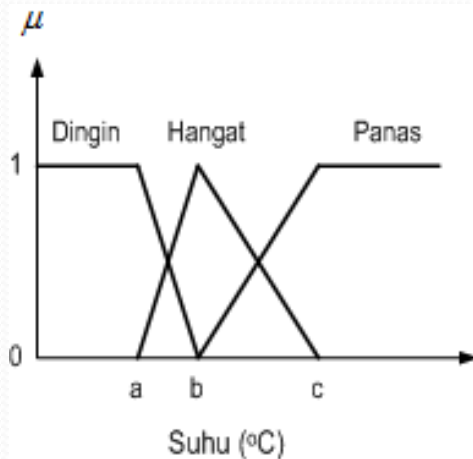
- Bagaimana cara menentukan **Fungsi Keanggotaan** dan/atau **Aturan** untuk *Fuzzy System* menggunakan EAs?
- Buat representasi **Individu**
- Buat **fungsi fitness-nya**
- Jelaskan proses yang terjadi pada EAs

# Apa yang di-evolusikan?

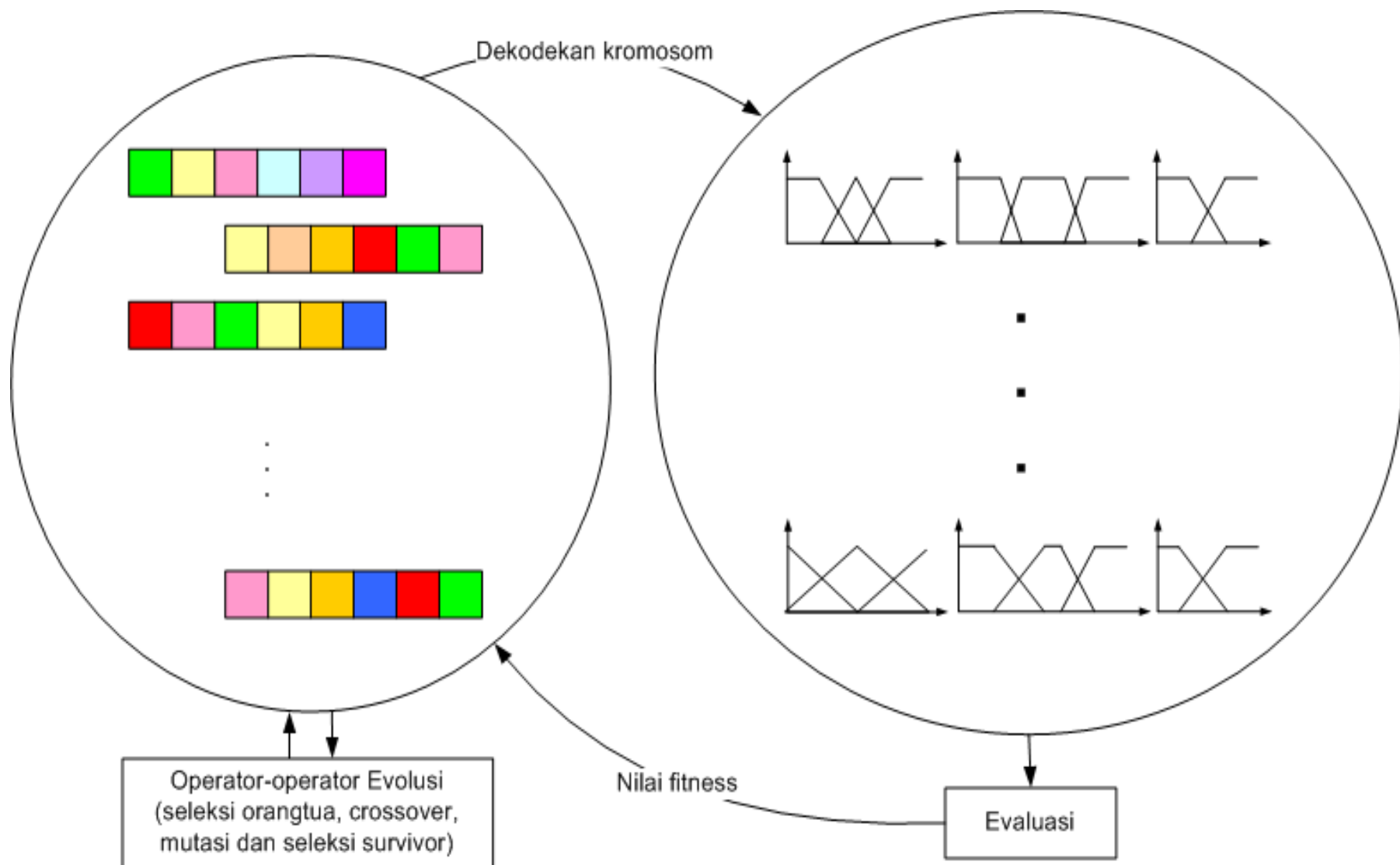
- Batas2 FK
- Jumlah dan Batas2 FK
- Bentuk, Jumlah, dan Batas2 FK
- Bentuk, Jumlah, dan Batas2 FK serta Rule

# Evolusi Batas2 FK

Asumsi: Jumlah dan Bentuk FK serta Fuzzy Rule sudah terdefinisi



1,351	2,117	3,015	0,986	1,289	3,201	4,712	5,377	12,078	54,098	80,986
a	b	c	d	e	f	g	h	i	j	k



## Evolusi Bentuk, Jumlah, dan batas2 FK serta Rule

- Kromosom sangat panjang
- Ruang solusi sangat besar
- Harus disesuaikan dengan kasus yang dihadapi

# *Sprinkler Control System*

Tanggal	Waktu	Suhu (°C)	Kelembaban Tanah (%)	Durasi Penyiraman (menit)
01-02-2006	08:00	20	16	55,7
01-02-2006	13:00	25	12	59,3
01-02-2006	18:00	16	30	5,6
02-02-2006	07:00	15	14	30,1
02-02-2006	12:00	27	10	43,4
02-02-2006	19:00	12	19	18,6
03-02-2006	06:30	19	12	22,1
03-02-2006	10:00	28	8	47,3
03-02-2006	13:30	35	6	76,4
03-02-2006	16:00	20	17	20,9

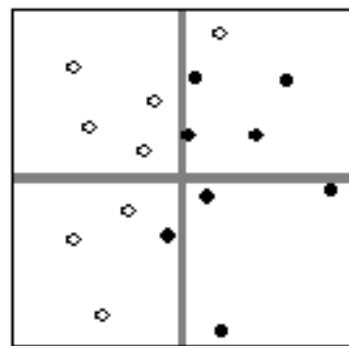
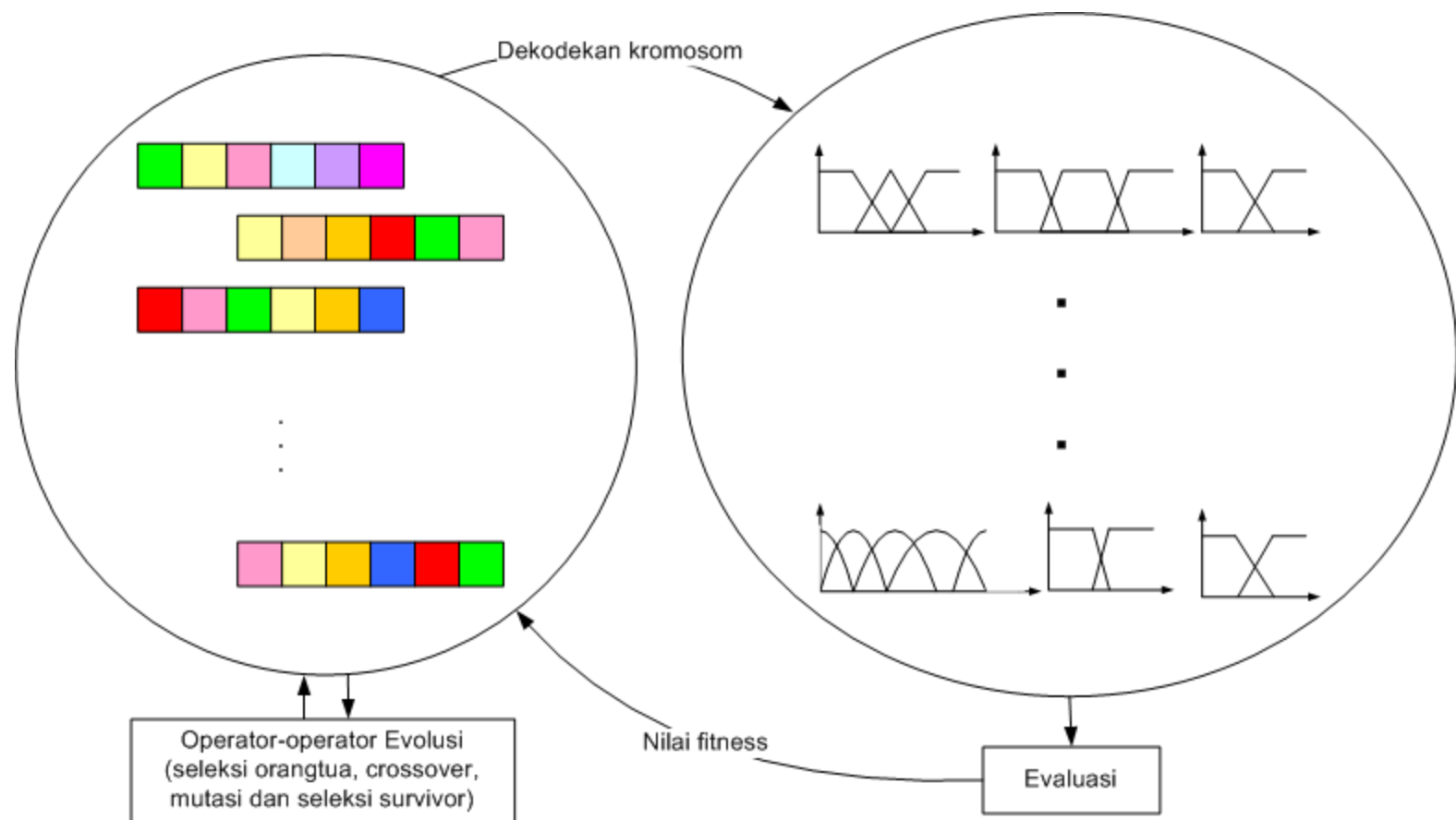
# Personalized Spam Filtering

<b>Task</b>	<b>File name</b>	<b>Data set size</b>	<b>Description</b>
Task A	task_a_labeled_train.tf	4000 emails	Labeled training emails.
	task_a_u00_eval.tf ... task_a_u02_eval.tf	2500 emails each	Unlabeled evaluation data: 3 inboxes from different users.
	task_a_labeled_tune.tf	4000 emails	Labeled training emails for parameter tuning. Feature representation corresponds only to file task_a_u00_tune.tf.
	task_a_u00_tune.tf	2500 emails	Labeled test emails of one user's inbox for parameter tuning. Feature representation corresponds only to file task_a_labeled_tune.tf.
Task B	task_b_labeled_train.tf	100 emails	Labeled training emails.
	task_b_u00_eval.tf ... task_b_u14_eval.tf	400 emails each	Unlabeled evaluation data: 15 inboxes from different users.
	task_b_labeled_tune.tf	100 emails	Labeled training emails for parameter tuning. Feature representation corresponds only to files task_b_u00_tune.tf and task_b_u01_tune.tf.
	task_b_u00_tune.tf task_b_u01_tune.tf	400 emails each	Labeled test emails from two user's inboxes for parameter tuning. Feature representation corresponds only to file task_b_labeled_tune.tf.

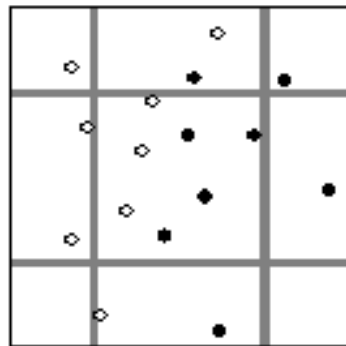


# Apa yang di-evolusikan?

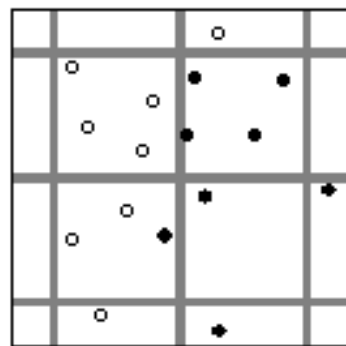
- Batas2 FK
- Jumlah dan Batas2 FK
- Bentuk, Jumlah, dan Batas2 FK
- Bentuk, Jumlah, dan Batas2 FK serta Rule



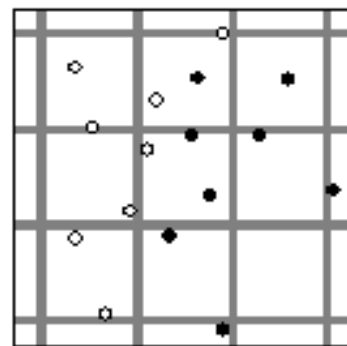
$K=2$



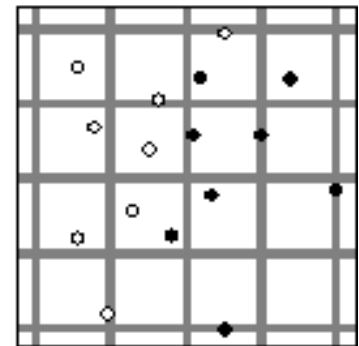
$K=3$



$K=4$



$K=5$



$K=6$

# *Personalized Spam Filtering*

- Bagaimana cara menentukan Fungsi Keanggotaan dan/atau Aturan untuk *Fuzzy System* menggunakan EAs?
- Buat representasi Individu
- Buat fungsi fitness-nya
- Jelaskan proses yang terjadi pada EAs. Apakah perlu dilakukan *preprocessing* data latih?

# Referensi

- [NEGo2] Negnevitsky, 2002, Hybrid intelligent systems: Evolutionary neural networks and fuzzy evolutionary systems, Pearson Education.
- [SUYo8] Suyanto, 2008, "Soft Computing: Membangun Mesin Ber-IQ Tinggi", Informatika, Bandung Indonesia. ISBN: 978-979-1153-49-2.
- [TETo1] Tettamanzi A., Tomassini M., "Soft Computing". Springer-Verlag Berlin Heidelberg, 2001. Printed in Germany.
- [MIT97] Mitchell M. Tom. 1997. "Machine Learning". McGraw-Hill International Editions. Printed in Singapore.