Partial Sampling Operator and Tree-Structural Distance for Multi-Objective GP

Makoto OHKI
Tottori University, Japan

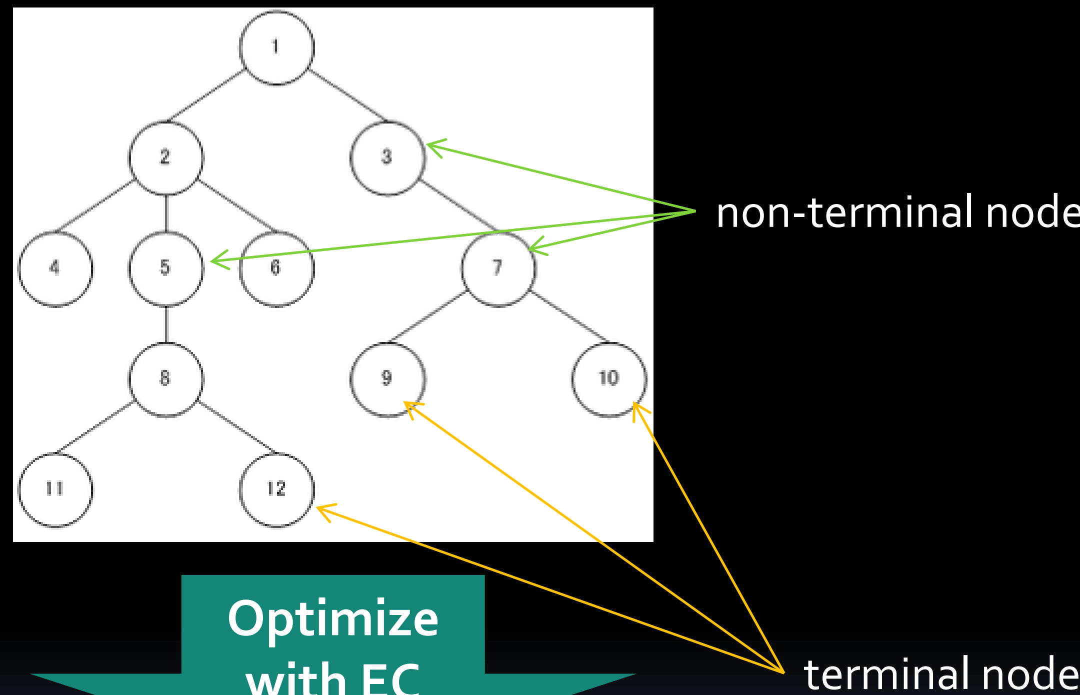
1. Introduction

Program Synthesis

Function Generation

Rule Discovery

- Target of these application
  → expressed by a tree structure.



Optimize with EC

Genetic Programming

1. Introduction

Effective Search ⇌ Bloat Control

* Schema Theory for GP [Holland 1992]
* Probabilistic Incremental Program Evolution [Salustowicz 1997]
* Depth Limitation [Langdon 1999]
* Size-Fair Model GP [Langdon 2000]
* Grammar-Guided GP [Ratle 2000]
* FREQT [Asai 2001]
* Subtree Swapping Crossover [Poli 2003]
* TAG3P [Hoai 2004]
* Tree Size Limitation [Ryan 2006]
* Stochastic Grammar-based GP [Ratle 2006]
* Semantic Building Blocks [McPhee 2008]

. . . . . .

Effective Search ⇌ Bloat Control

* Schema Theory for GP [Holland 1992]
* Probabilistic Incremental Program Evolution [Salustowicz 1997]
* Depth Limitation [Langdon 1999]
* Size-Fair Model GP [Langdon 2000]
* Grammar-Guided GP [Ratle 2000]
* FREQT [Asai 2002]
* Subtree Swapping Crossover [Poli 2003]
* TAG3P [Hoai 2004]
* Tree Size Limitation [Ryan 2006]
* Stochastic Grammar-based GP [Ratle 2006]
* Semantic Building Blocks [McPhee 2008]

. . . . . .

1. Introduction

Effective Search ⇌ Bloat Control

* Schema Theory for GP [Holland 1992]
* Probabilistic Incremental Program Evolution [Salustowicz 1997]
* Depth Limitation [Langdon 1999]
* Size-Fair Model GP [Langdon 2000]
* Grammar-Guided GP [Ratle 2000]
* FREQT [Asai 2001]
* Subtree Swapping Crossover [Poli 2003]
* TAG3P [Hoai 2004]
* Tree Size Limitation [Ryan 2006]
* Stochastic Grammar-based GP [Ratle 2006]
* Semantic Building Blocks [McPhee 2008]

. . . . . .

1. Introduction

Effective Search ⇌ Bloat Control

* Schema Theory for GP [Holland 1992]
* Probabilistic Incremental Program Evolution [Salustowicz 1997]
* Depth Limitation [Langdon 1999]
* Size-Fair Model GP [Langdon 2000]
* Grammar-Guided GP [Ratle 2000]
* FREQT [Asai 2001]
* Subtree Swapping Crossover [Poli 2003]
* TAG3P [Hoai 2004]
* Tree Size Limitation [Ryan 2006]
* Stochastic Grammar-based GP [Ratle 2006]
* Semantic Building Blocks [McPhee 2008]

. . . . . .

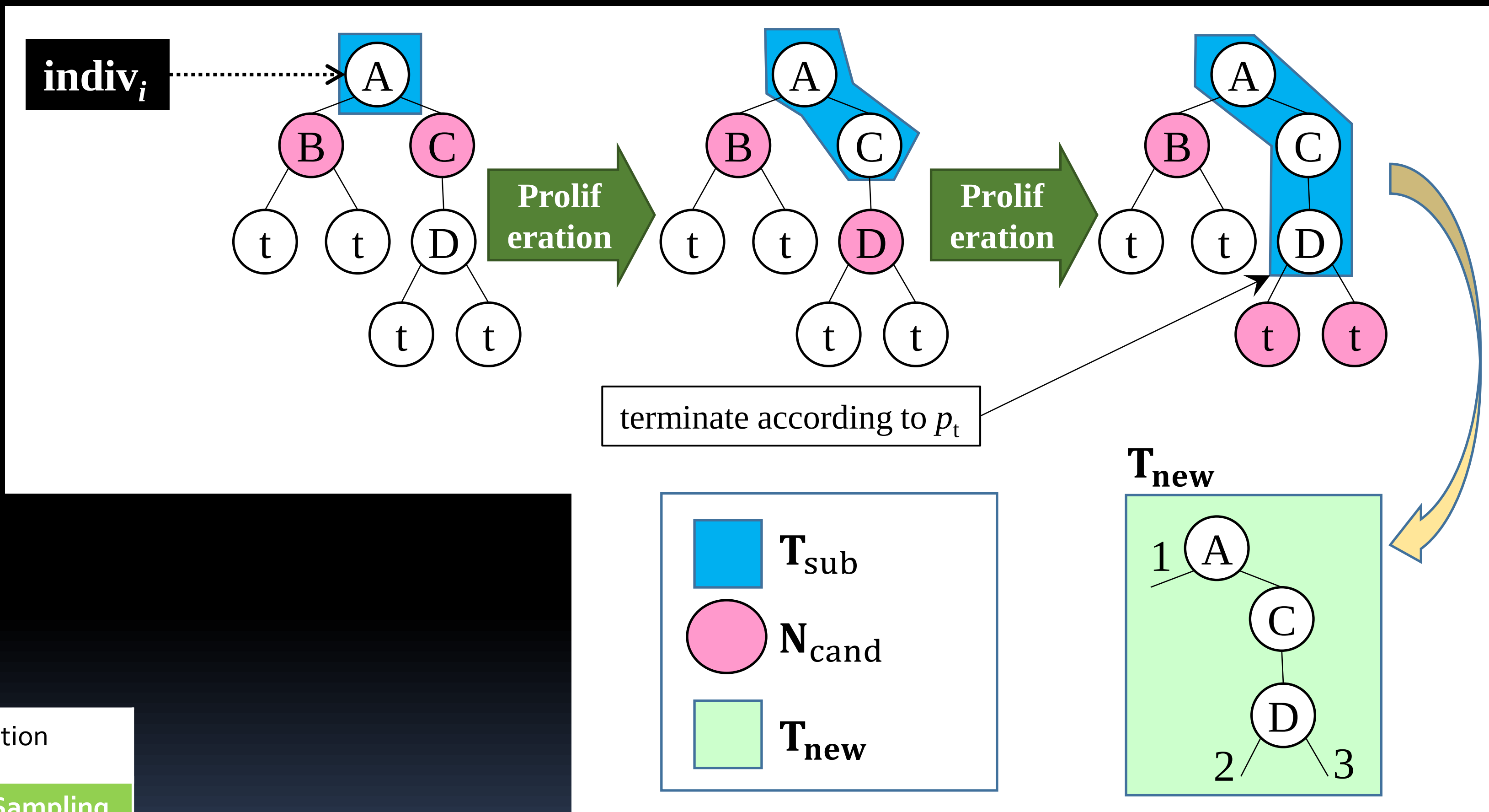
1. Introduction

In this paper,

- Partial Sampling (PS) operator instead of Crossover and Mutation
- A technique of Multi-Objective GP by applying NSGA-II
  - index of goodness of the tree
  - the Size of the tree
  - tree position in the population by Tree-Structural Distance (TSD)
- Apply TSD instead of Crowding Distance (CD) of NSGA-II
- Double Spiral Problem for verification

2. Partial Sampling Operator for Mating

- Proliferation in Partial Sampling (PS) Operator



2. Partial Sampling Operator for Mating

- Proliferation Termination Probability p_t

$$p_t^0 = \frac{1}{\text{AverageSize}\ \mathbf{R}^g},$$

$$p_t^{g+1} = \frac{\dfrac{1}{\text{Succeed}\ \mathbf{P}^g} - p_t^0}{\dfrac{1}{\text{Succeed}\ \mathbf{R}^g} - p_t^0}\ p_t^g - p_t^0\ + p_t^0,$$

$\mathbf{R}^g$: population at g-th generation

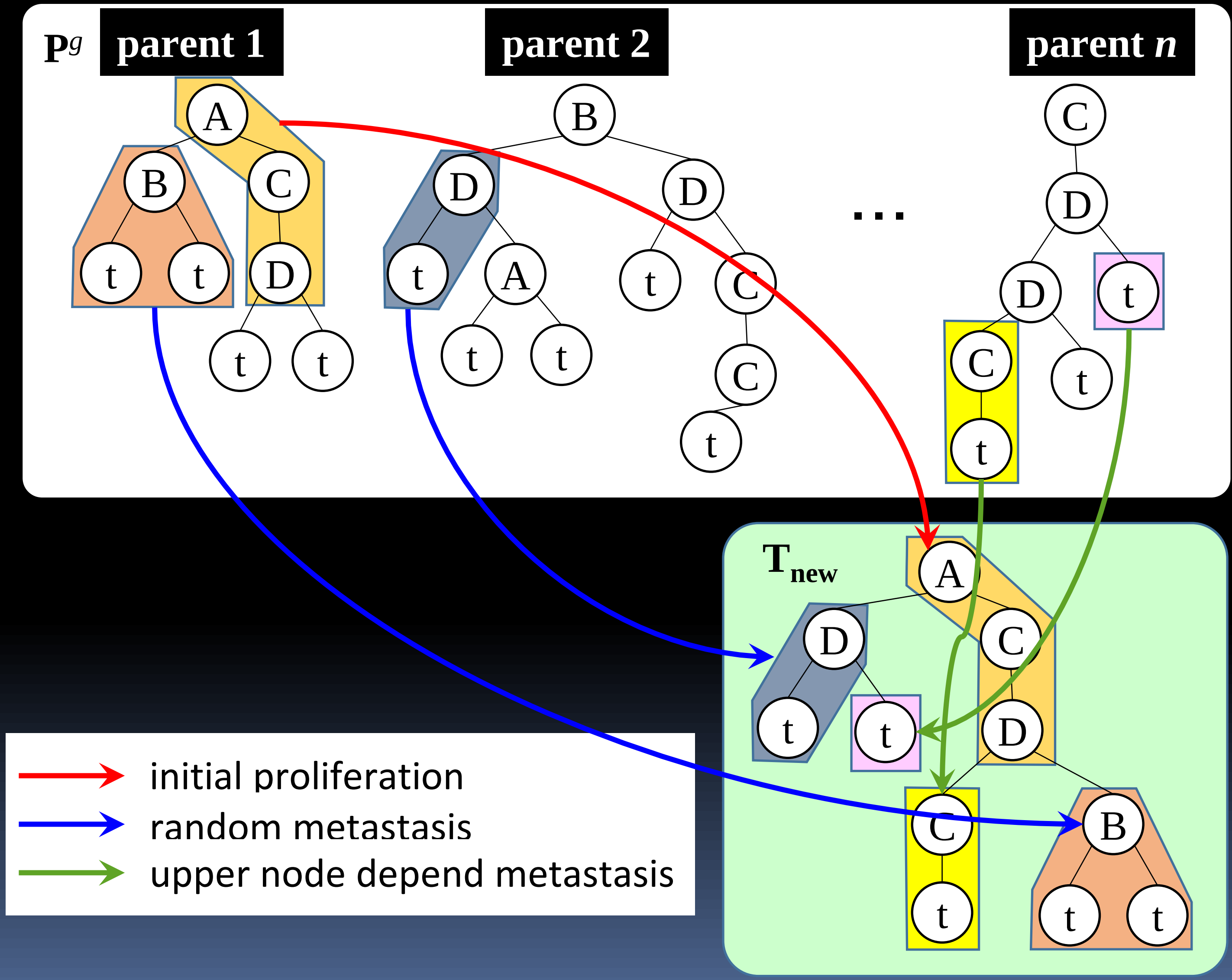
$\mathbf{P}^g$: parent set at g-th generation

AverageSize $\bullet$: average size of tree structure

Secceed $\bullet$: average size of partial tree structure of set succeeded from previous generation

2. Partial Sampling Operator for Mating

- 2 kinds of metastasis



3. Multi-Objective GP with Tree-Structural Distance

- 3 Objective Functions

① objective function according to Goodness of the tree structure

$$h_1(\mathrm{indiv}_i) = \mathrm{performance}(\mathrm{root}_i)$$

② objective function according to the size of the tree structure

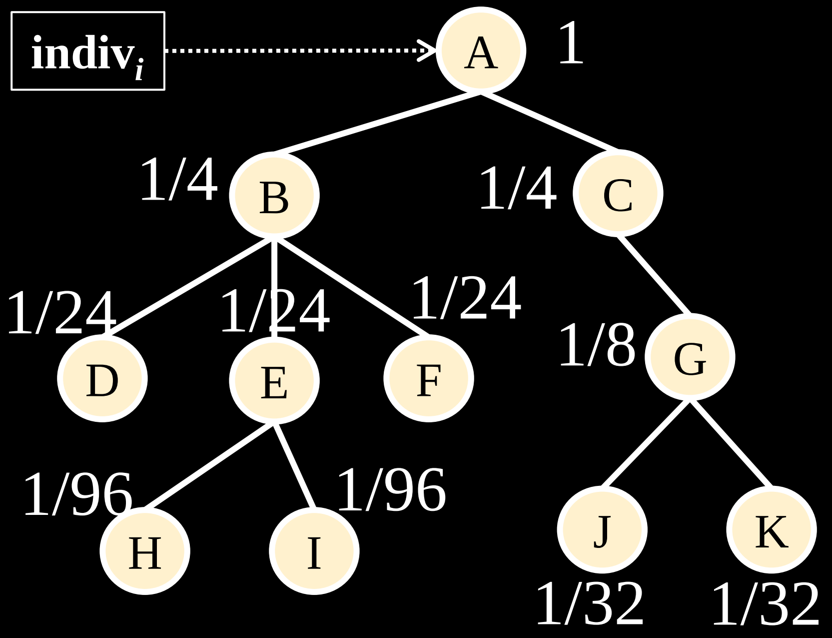
$$h_2(\mathrm{indiv}_i) = \frac{1}{\mathrm{size}(\mathrm{root}_i)}$$

③ objective function according to average of TSD in the population

$$h_3(\mathrm{indiv}_i) = \frac{1}{N_{\mathrm{pop}}} \sum_{k=1}^{N_{\mathrm{pop}}} \mathrm{TSD}\ \mathrm{indiv}_i, \mathrm{indiv}_k$$

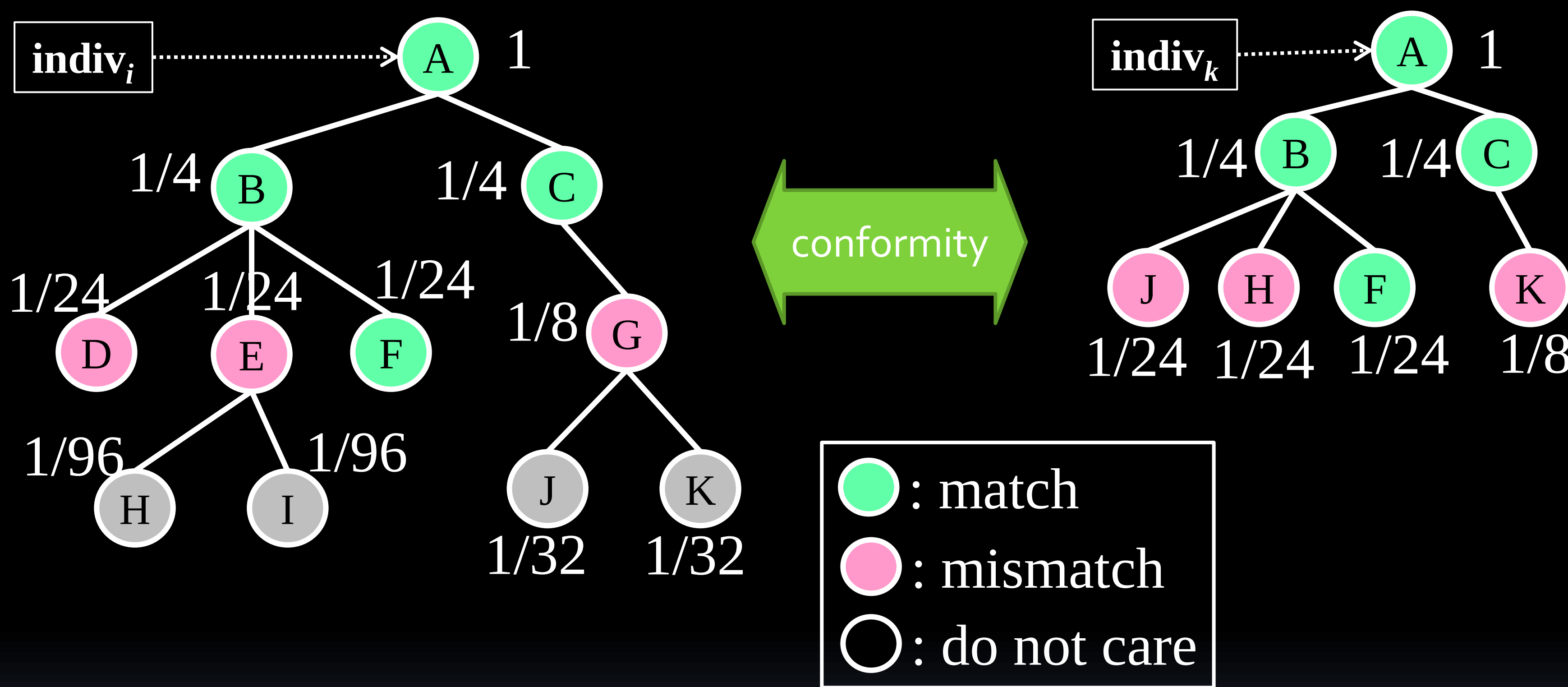
3. Multi-Objective GP with Tree-Structural Distance

- Tree-Structural Distance (TSD)



3. Multi-Objective GP with Tree-Structural Distance

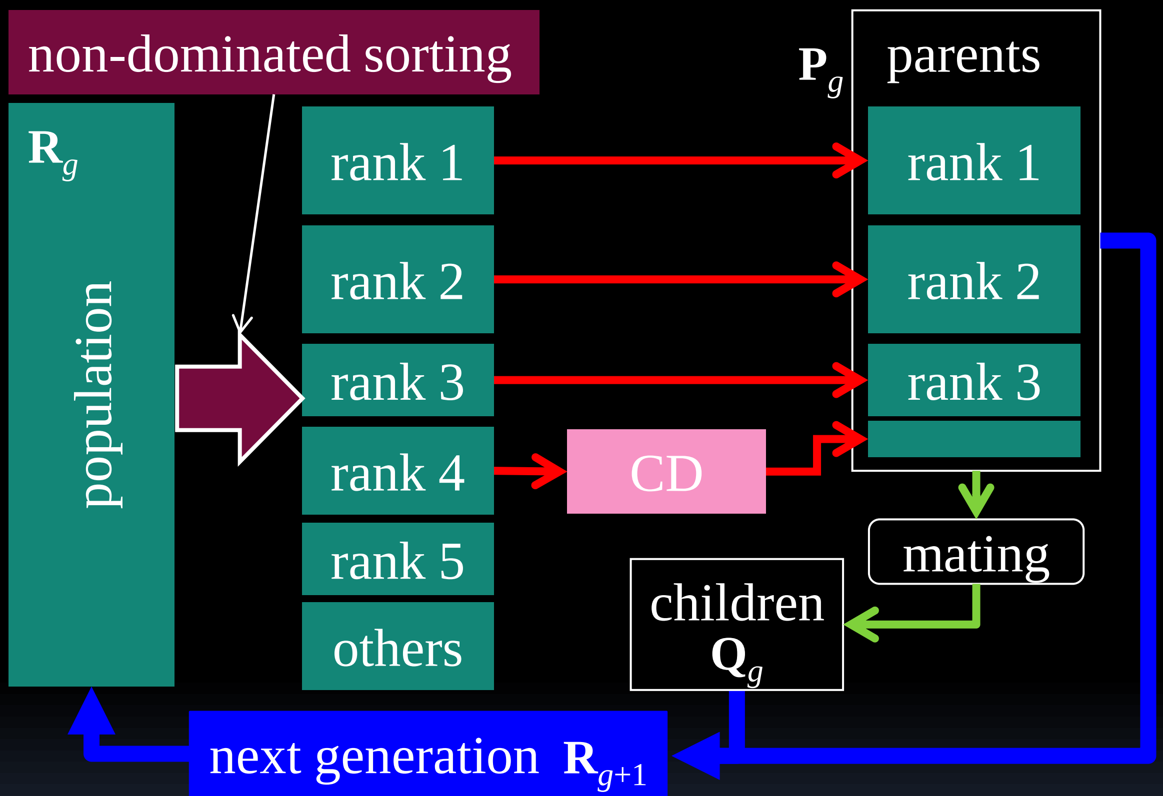
- Tree-Structural Distance (TSD)



$$\mathrm{TSD}(\mathrm{root}_i, \mathrm{root}_k) = \frac{1}{24} + \frac{1}{24} + \frac{1}{8} = \frac{5}{24}$$

3. Multi-Objective GP with Tree-Structural Distance

- NSGA-II (conventional)

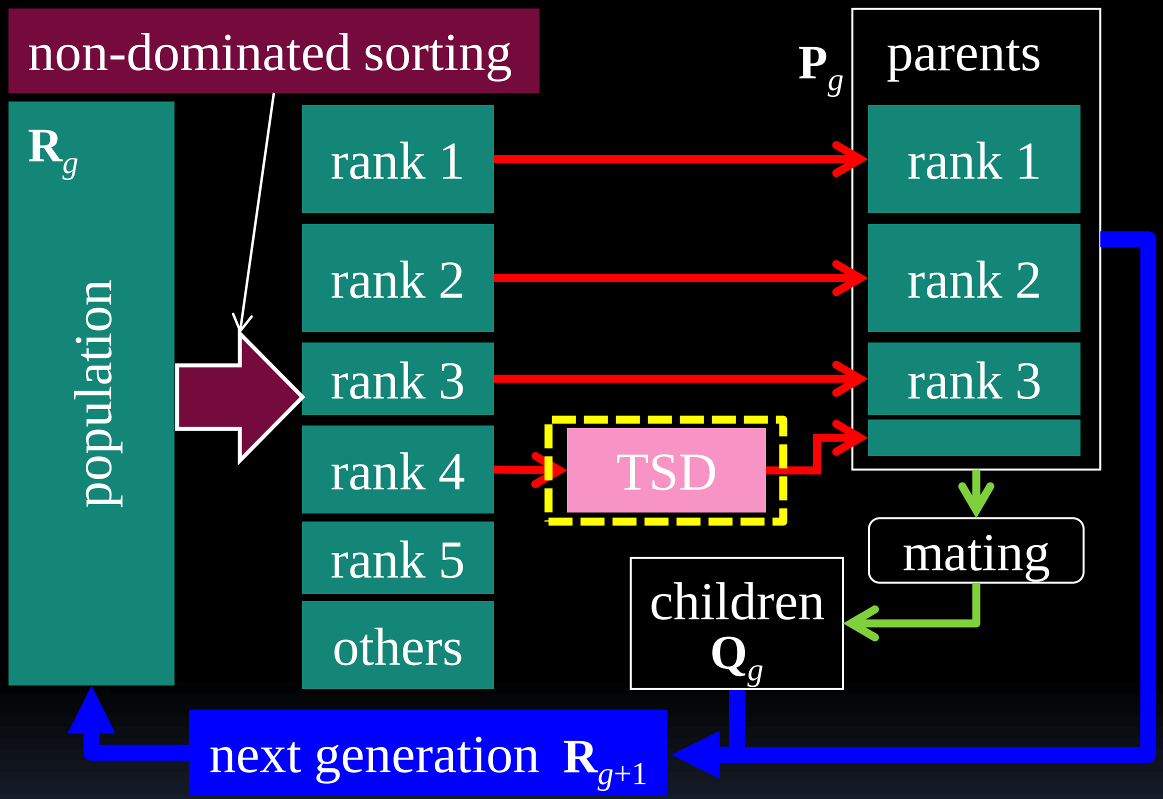


R^g : population

P^g : parents

3. Multi-Objective GP with Tree-Structural Distance

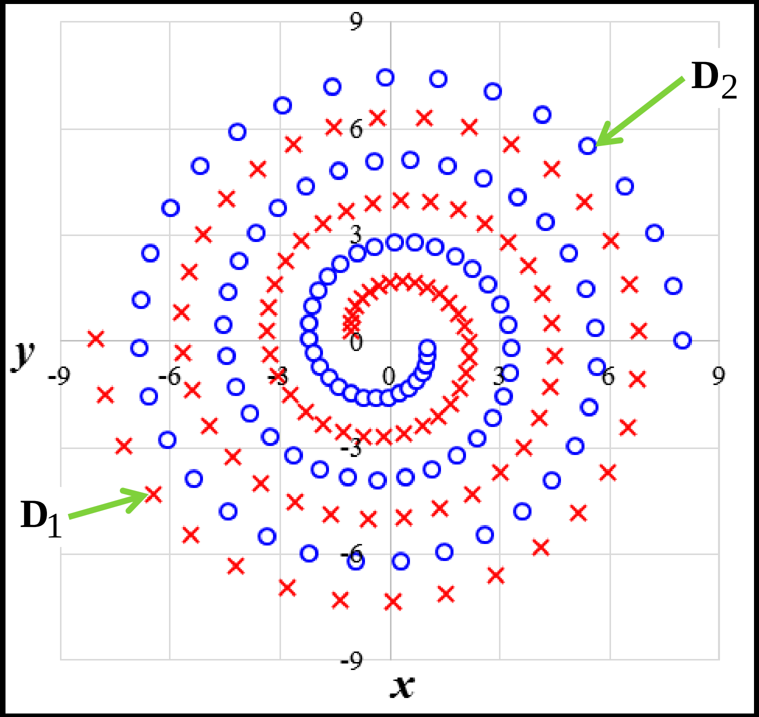
- NSGA-II with TSD instead of CD



R^g : population

P^g : parents

4. Verification by Double Spiral Problem



$$\begin{cases} f(x,y) > 0 \Leftrightarrow (x,y) \in \mathbf{D}_1 \\ f(x,y) < 0 \Leftrightarrow (x,y) \in \mathbf{D}_2 \\ f(x,y) = 0 \Leftrightarrow \text{FALSE} \end{cases}$$

difficult even by the neural network.

GP

- ○ non-terminal node $\in$ $+, -, *, \div, \sin, \cos, \tan, \text{ifltz}$
- ○ terminal node $\in$ $x, y, \text{constant}$

$$\text{ifltz}(a,b,c) \triangleq \text{if } a < 0 \text{ then } b \text{ else } c = \begin{cases} b & (a < 0) \\ c & (\text{otherwise}) \end{cases}$$

4. Verification by Double Spiral Problem

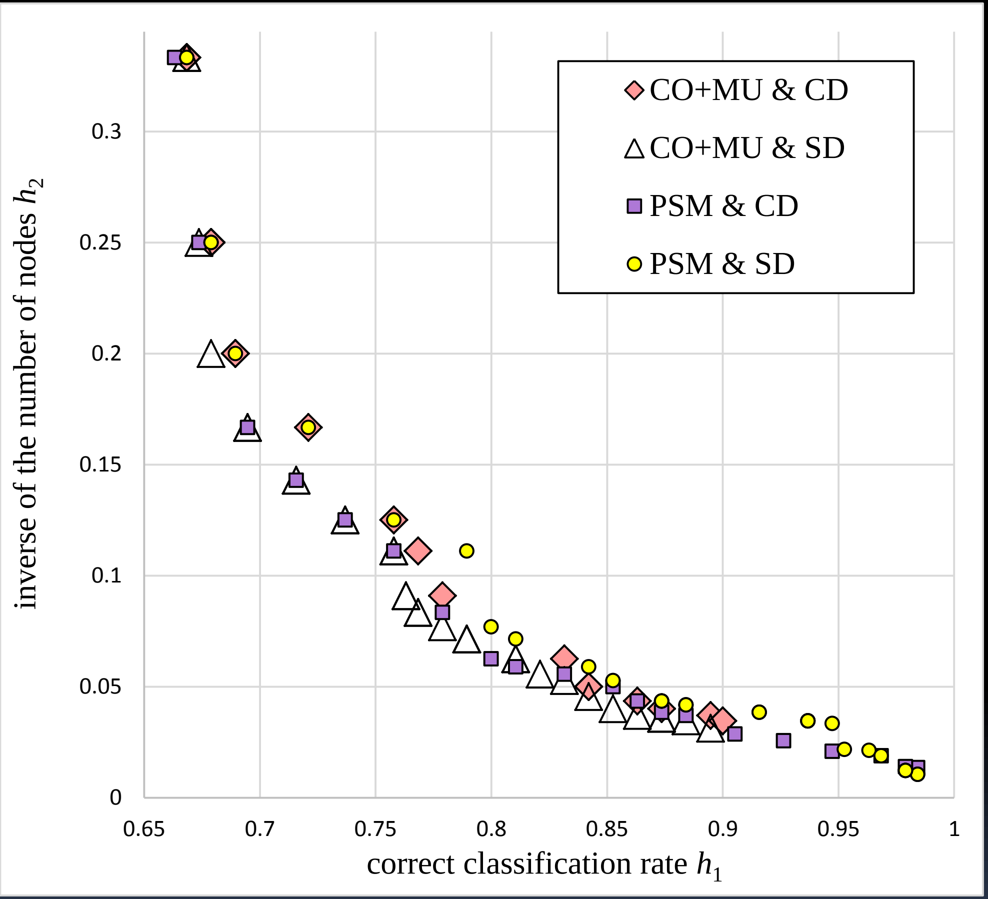
- Objective function h_1 according to the goodness of tree

$$h_1(\mathrm{indiv}_i) = \mathrm{performance}(\mathrm{root}_i) = \frac{1}{|\mathbf{D}_1 \cup \mathbf{D}_2|} \sum_{k=1}^{|\mathbf{D}_1 \cup \mathbf{D}_2|} g(x_k, y_k)$$

$$g(x,y) = \begin{cases} 1 & f(x,y) > 0 \wedge (x,y) \in \mathbf{D}_1 \ , \\ 0 & f(x,y) > 0 \wedge (x,y) \in \mathbf{D}_2 \ , \\ 1 & f(x,y) < 0 \wedge (x,y) \in \mathbf{D}_2 \ , \\ 0 & f(x,y) < 0 \wedge (x,y) \in \mathbf{D}_1 \ , \\ 0 & f(x,y) = 0 \end{cases}$$

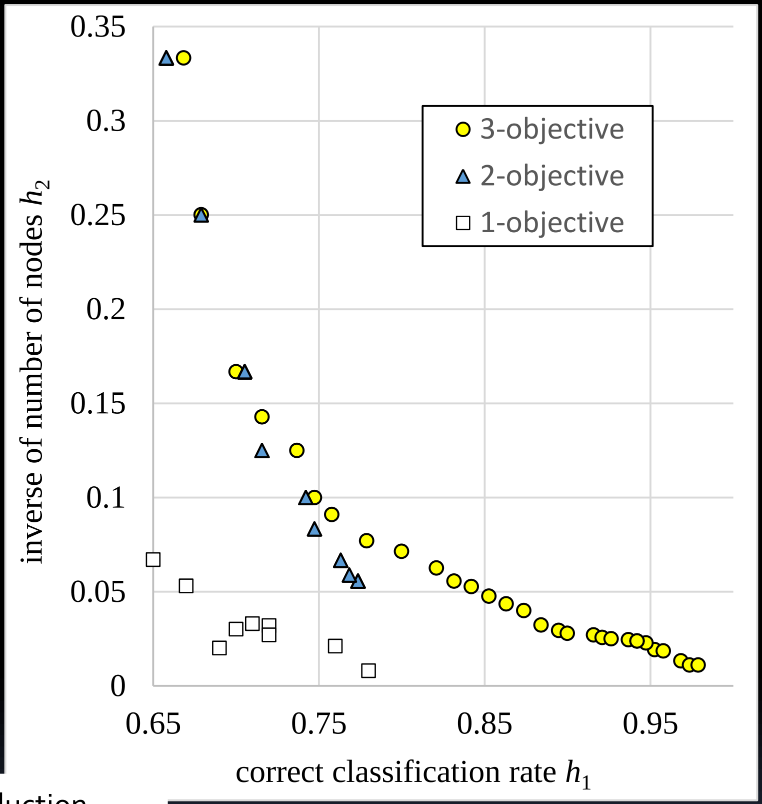
4. Verification by Double Spiral Problem

- Final Solution Distribution

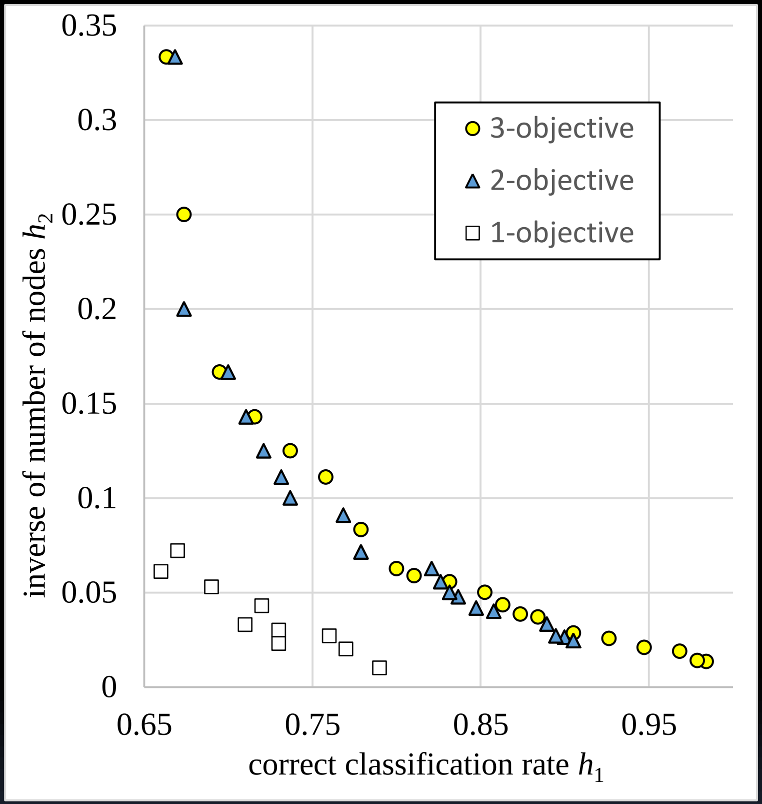


4. Verification by Double Spiral Problem

- Comparison among 3-Objective, 2-Objective, 1-Objective GPs



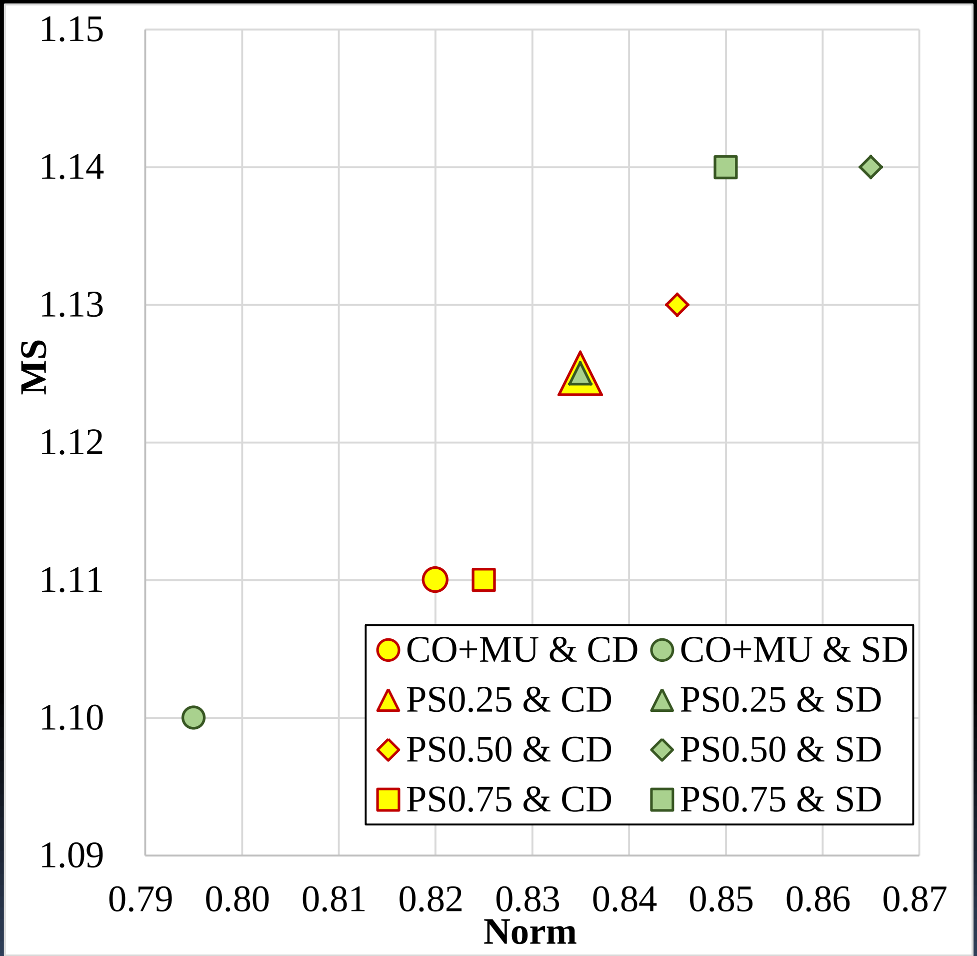
with SD



with CD

4. Verification by Double Spiral Problem

- Comparison of results on MS-Norm plane



5. Conclusion

In this paper,

- Multi-Objective GP
- In addition to goodness of the tree, 2 objective functions
  - tree size
  - Tree-Structural Distance (TSD)
- Partial Sampling (PS) for mating

- Double Spiral Problem for verification.

- The proposed technique (PS + TSD → NSGA-II) is effective.

5. Conclusion

In the future,

- Enhance the capability of numerical optimization
- Ranking Selection technique harmonizing CD and TSD
- Mechanism to forcibly exit from PS

Thank you very much!

Ask me simply, even if you have.

- MS and Norm

degree of spread of $\mathcal{FFS}$

$$\mathrm{MS} = \sqrt{\sum_{i=1}^{m} \left(\max_{j=1}^{|\mathcal{FFS}|} f_i(\mathbf{x}_j) - \min_{j=1}^{|\mathcal{FFS}|} f_i(\mathbf{x}_j) \right)^2}$$

degree of convergence to $\mathcal{POS}$

$$\mathrm{Norm} = \frac{1}{|\mathcal{FFS}|} \sum_{j=1}^{|\mathcal{FFS}|} \sqrt{\sum_{i=1}^{m} f_i(\mathbf{x}_j)^2}$$

1. Introduction

- Program Synthesis [David2017]

```
type list = Nil | Cons of int * list
let rec even x =
   match x with
     | Nil -> Nil
     | Cons(u, Nil) -> Cons(u, Nil)
     | Cons(u, Cons(_, us))
          -> Cons(u, even us)
let rec sum x =
   match x with
     | Nil -> 0
     | Cons(u, us) -> u + sum us
let rec sum_even = ??
let main x = assert (sum (even x) = sum_even x)
```

solution

```
let rec sum_even x =
   match x with
     | Nil -> 0
     | Cons (u, Nil) -> u
     | Cons (u, Cons(_, us)) -> u + sum_even us
```

1. Introduction

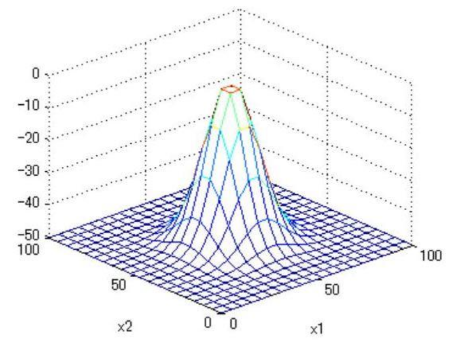
Applications of Genetic Programming (GP)

- Program Synthesis
- Function Generation
- Rule Set Discovery

· · · · · · ·

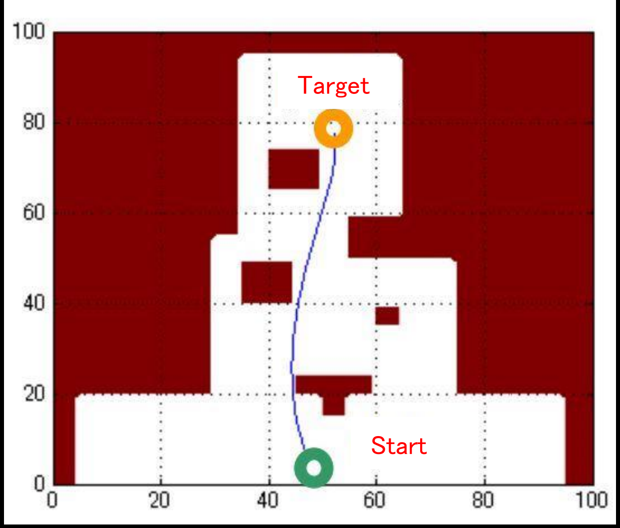
1. Introduction

- Function Generation [Jamali2017]

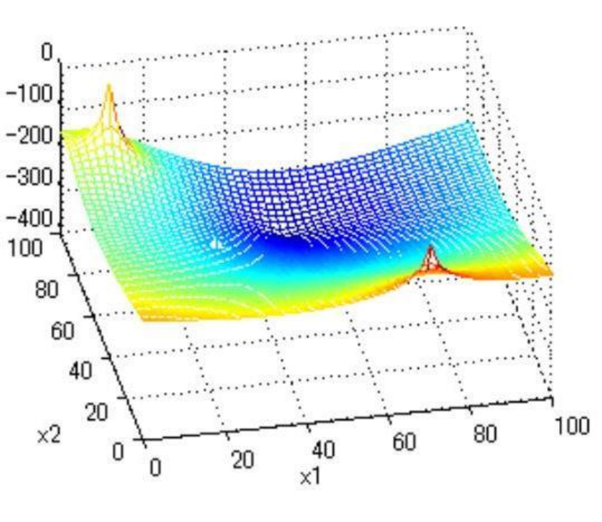


$$U(x_p) = -Ce^{-\frac{(x_p - x_g)^2}{r^2}}$$

Potential Function

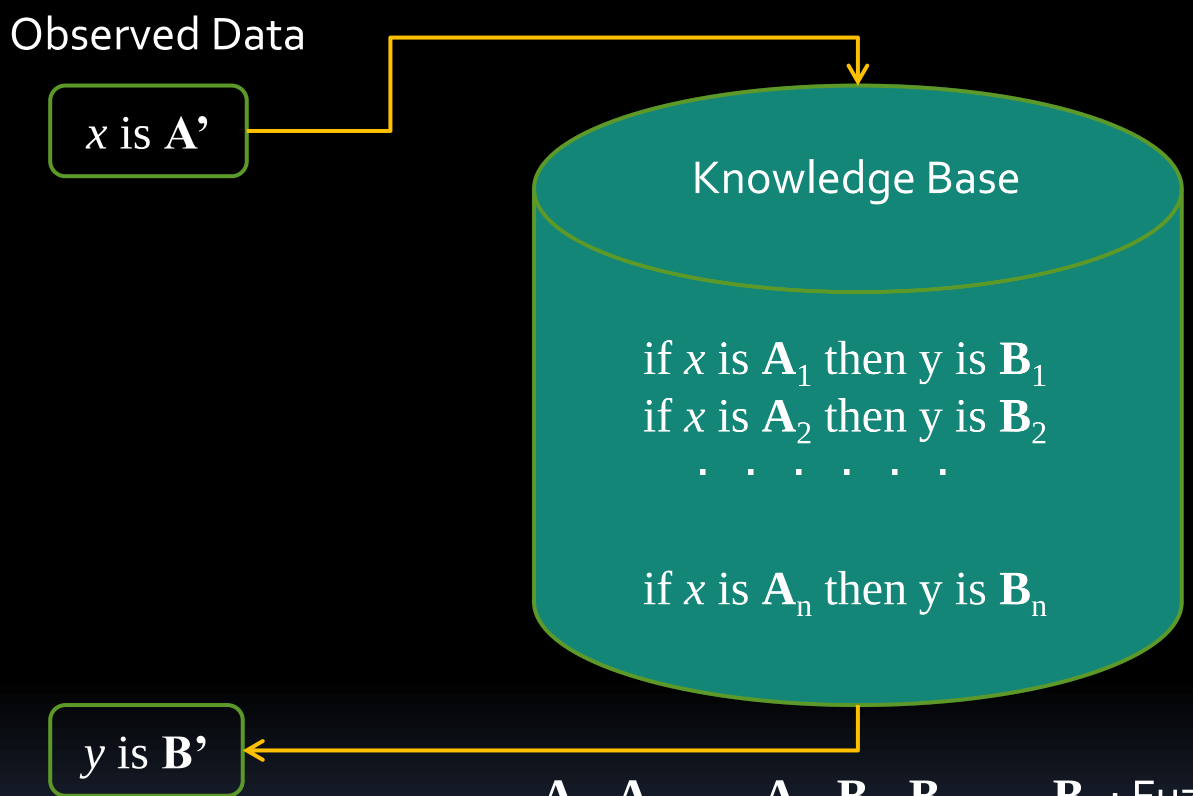


solution



1. Introduction

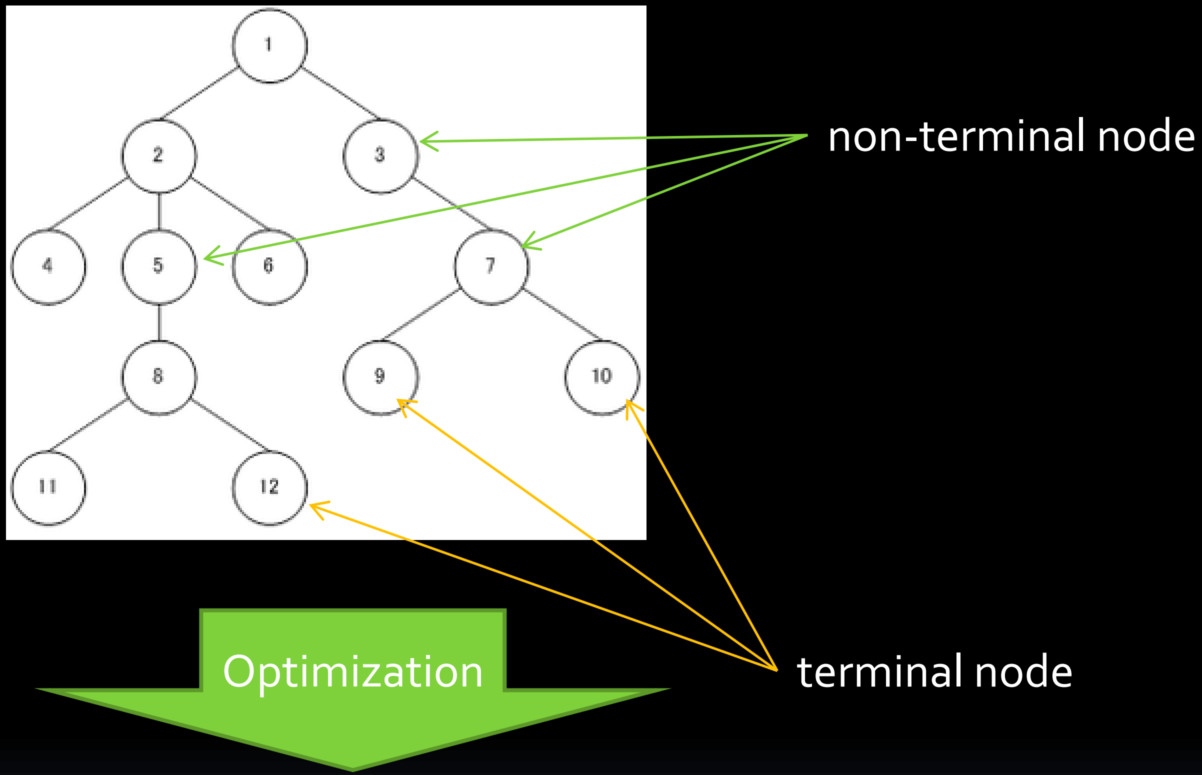
- Rule Set Discovery [Ohmoto2013]



$\mathbf{A}_1, \mathbf{A}_1, \ldots, \mathbf{A}_n, \mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n$: Fuzzy Sets

1. Introduction

- They can be expressed by a tree structure data.



Genetic Programming: GP

1. Introduction

- Intron