Function Finding and the Creation of Numerical Constants in GEP

Cândida Ferreira

AIM

Analyse the usefulness of numerical constants in evolutionary computation

Problem Suite (1)

SEQUENCE INDUCTION

- Computer generated
- Integer constants
- 1 variable

Problem Suite (2)

V FUNCTION

- Computer generated
- Rational constants
- 1 variable

PROBLEM SUITE (3)

WOLFER SUNSPOTS

- Real-world problem
- Rational constants
- 10 variables

RESULTS: Performance (1)

SEQUENCE INDUCTION

	With constants	Without constants
Number of runs	100	100
Number of generations	100	100
Population size	100	100
Average best-of-run fitness	179.827	197.232
Average best-of-run R-square	0.977612	0.999345
Success rate	16%	81%

RESULTS: Performance (2)

V FUNCTION

	With constants	Without constants
Number of runs	100	100
Number of generations	5000	5000
Population size	100	100
Average best-of-run fitness	1914.8	1931.84
Average best-of-run R-square	0.957255	0.995340

RESULTS: Performance (3)

WOLFER SUNSPOTS

	With constants	Without constants
Number of runs	100	100
Number of generations	5000	5000
Population size	100	100
Average best-of-run fitness	86215.27	89033.29
Average best-of-run R-square	0.713365	0.811863

RESULTS: Best evolved models (1)

SEQUENCE INDUCTION

With constants

$$y = 5a^4 + 4a^3 + 3a^2 + 2a + 1$$

R-square: 1

Without constants

$$y = 5a^4 + 4a^3 + 3a^2 + 2a + 1$$

R-square: 1

RESULTS: Best evolved models (2)

V FUNCTION

With constants

$$y = \left[\ln\left(0.99782a^2\right)\right] + \left[10^{\sin(1.27278a)}\right] + \left[10^{0.929a}\right] + \left[0.77631 - 2.80112a^3\right] + \left[2.45714 + e^{0.981a} + e^a\right]$$

R-square: 0.9999313

Without constants

$$y = \left[\ln\left(2a^2\right) + 10^{\sin a}\right] + \left[2a + \sin a + a^2\right] + \left[\cos(\cos(2a)) + e^{a^2}\right] + \left[e^{\sin a}\right] + \left[1 + e^a + e^{a^2}\right]$$

R-square: 0.99997001

RESULTS: Best evolved models (3)

WOLFER SUNSPOTS

With constants

$$y = \frac{2j^2}{h+i+j} + \frac{a+b+g}{0.995+0.847c+e} + \frac{1.903j+j^2}{i^2+j}$$

R-square: 0.833714

Without constants

$$y = j + \frac{d-i+3j}{b+e} + \frac{d+bj-ij}{a+2i}$$

R-square: 0.882831

CONCLUSIONS

The use of numerical constants results in:

- worse performance
- more complex implementation
- worse evolution
- more CPU time