Meme as Building Block for Evolutionary Optimization
A Machine & Transfer Learning Paradigm for Search

presented by

Yew-Soon, Ong
Director, Center for Computational Intelligence
Director, SIMTECH-NTU Joint Lab on Complex Systems
Principal Investigator, NTU-Rolls Royce Corporate Joint Lab
School of Computer Engineering, Nanyang Technological University
Nanyang Avenue, Singapore
Presentation Outline

- Introduction and Motivation
- Meme as building Block for Evolutionary Optimization
- Memetic Search with Inter-Domain Learning
- Conclusion
Evolutionary Algorithm

Algorithm 1: Evolutionary Algorithm

1 Begin
2 \( k := 0 \) /* Initialize the evaluation counter. */
3 Initialize and evaluate \([P(k)]\); /* Create an initial population. */
4 while Stopping conditions are not satisfied do
5 \[ P'(k) := \text{Reproduction}[P(k)] \] /* Apply cross and mutation operators. */
6 Evaluate \([P'(k)]\)
7 \( P(k + 1) := \text{select}[P'(k), P(k)]; \) /* Create an new population. */
8 \( k := k + 1; \) /* Increase the evaluation counter. */
9 End

- Stochastic algorithm
- Successfully be applied for solving complex engineering optimization problems
- Term ‘Memetic Algorithm’ was introduced by Moscato in 1989.
- Hybrids of population-based search and individual learning procedure.
- Meme: Local search, individual learning, local refinement process.

**Algorithm 2: A Canonical Memetic Algorithm**

```plaintext
1 Begin:
2 Initialize: Generate an initial GA population.
3 while Stopping conditions are not satisfied do
4     Evaluate all individuals in the population
5     for each individual in the population do
6         Proceed with local improvement and replace the genotype and/or phenotype in the population with the improved solution in the spirit of Lamarckian or baldwinian Learning
7     Apply standard GA operators to create a new population; i.e., crossover, mutation and selection
8 End
```
Evolutionary, Memetic Algorithm

Start a search from scratch or Ground Zero Knowledge State

Capability does not grow or evolve along with problem solved

or

experiences!!!
Memetic Computation

Shall we look at Memes as Truly Building Blocks that culminate as recurring information patterns & latent structures, i.e. Higher-Order Representation?

“meme is a unit of information residing in the brain and is the replicator in human cultural Evolution”

From basic patterns or structures, complex higher level structures can then be evolved and discovered efficiently.
The Term ‘Memetic Computing’ founded in 2007

Founding Chair: Yew-Soon, Ong

Emergent Technologies Task Force on Memetic Computing

'Semetic Computing' referred by Thomson Scientific's Essential Science Indicators as an Emerging Front, August 2008.

Some of the Organized Events

Special Session on Memetic Computing, IEEE World Congress on Computational Intelligence, WCCI 2014, July 6-11, Beijing, China, Organizers: Zexuan Zhu, Wenyin Gong, Maoguo Gong and Yew-Soon Ong.


Dawkins in 1976, Book “The Selfish Gene”, Chapter 11:

“meme is a unit of information residing in the brain and is the replicator in human cultural Evolution”

New Science of Memetics represents Mind-Universe analog to genetics in cultural evolution, stretching from anthropology, biology, cognition, psychology, sociology and socio-biology.

Greek Origin:
‘To imitate’

“Mimeme” → “Meme”

Mono-syllabus French:
‘Memory’
Roles of Gene, Meme & associations

Gene → instructions for making proteins

Meme → instructions for carrying out behaviour & making of artifacts

Building Blocks: Gene in Genetics → Meme in Memetics

Book “The Meme Machine” → by Blackmore
“reaffirmed meme definition as information copied from one person to other person.”

• Discusses theory of "memetic selection"- survival of the fittest among competitive ideas down through generations

• Delved on ‘possibility of meme developing into a proper hypothesis of the human mind’
What defines and constitutes a Meme?

Multi-facet views of “meme” as:

- information restricted to the brain and expressed in forms of behaviours
- contagious information pattern that replicates by parasitically infecting human minds
- constellation of activated neuronal synapses in memory & information encoded in neural structure
- memory item, or portion of an organism's neurally-stored information
- genotype/memotype as mental representation, and phenotype/sociotype as behaviour
Memetic Computation in Nature-Inspired Optimization
A Bottom-Up Search Paradigm
(Meme as Higher-Order Representation of Gene)

- Define the memes or building blocks
- Mining, Discovery & Learning of Memes
- Meme Transmission, Variations and Selection


"meme is a unit of information residing in the brain and is the replicator in human cultural Evolution"
Towards Self-Assembly of Meme like in Nature

Cluster and Small Biochemical systems

Super-molecules

Prokaryotic and Eukaryotic Cells

Atoms

Energy Landscape
Motivation

- Problems seldom exist in isolation & hence Human **DO NOT** search from Scratch
  - Useful information exist in past problems & when properly harnessed are effective for future problem-solving.

Some Related Work in Routing/Scheduling Problem

- Cased Injected Genetic Algorithm (Louis et al., 2004)
- High-quality scheduling reuse in TSP (Cunningham et al., 1997)
- **Drawback:** if the given new problem happens to differ in problem size, structure, problem domain, etc., what have been previously memorized from past problems cannot be directly injected into the search for reuse.
Meta-Learning: Higher Order Learning in Search
(Discovery and Generalization of Memes as Recurrence of Patterns or Latent Structures)
Even and odd parity problems, Standard benchmarks for genetic programming (Koza 1992)

Even-2-parity function is simply the XOR function, i.e., a composition of the terminal and function set in one simple possible configuration:  
\[ a \text{ XOR } b = (a \text{ OR } b) \text{ AND } (a \text{ NAND } b) \]

Even-3-parity function using same primitives is more difficult, but follows a similar pattern, illustrated below and in Figure 7:
\[ \text{XOR} (a, b, c) = (((a \text{ OR } b) \text{ AND } (a \text{ NAND } b)) \text{ OR } c) \text{ AND } (((a \text{ OR } b) \text{ AND } (a \text{ NAND } b)) \text{ NAND } c) \]

**Meta-Learning XOR3 to XOR-N**

**Evolution**

The probability that the system will transition from XOR2 to XOR3 to XOR-N is calculated using only the mutation step.

Memetic Computing
Meta-Learning: Higher Order Learning
(Discovery and Generalization of Memes as Recurrence of Patterns or Latent Structures)
A meta-learning system is composed of primary components:

- **Optimizer** - specialize solutions on given problem instance.
- **Memory** - stores & generalized previous solutions encountered & passes selected solution(s) or memes to optimizer.
- **Selection** - takes features of given problem, & performs mapping to solutions in memory.
- **Generalization** - compares resultant solution with solutions in memory, adds/modifies solution.

A Realization of Meta-Learning for Real World Complex Problem

Meme as building Block for Evolutionary Optimization

\[ P_{\text{new}} = \{ p^j | j = n + 1, \ldots, \infty \} \]

\[ P_{\text{old}} = \{ p^i | i = 1, \ldots, n \} \]

\[ S_{\text{new}}^j = \{ s^g | g = 1, \ldots, \text{PopSize} \} \]

(a) Conventional evolutionary solver (ES)

(b) Proposed culture-inspired operators
Meme as Building Blocks for Evolutionary Optimization

**Learning Operator**
- Capture memes from past problem solving experiences
- Proceeds in an incremental manner

**Selection Operator**
- Identify the suitable building blocks or meme(s) from the society of memes to operate on future unseen problems

**Variation Operator**
- The meme variation forms the intrinsic innovation tendency of the cultural evolution.

**Imitation Operator**
- Memes that are learned from past problem solving experiences replicates by means of imitation and used to enhance future evolutionary search on newly encountered problems.

A Case Study

- **Routing Problem Domains (CVRP and CARP)**
  - Capacitated vehicle routing and capacitated arc routing
  - Servicing a set of customer (represented as vertices in CVRP and streets or edges in CARP)
  - Using a fleet of capacity constrained vehicles
  - Vehicles located at the central depot
  - Minimize the total routing cost involved
  - Combinatorial optimization problem
  - NP-Hard

These consist of problems of diverse properties in terms of vertices size, graph topologies, etc.,

which learning from problems and across domains cannot be easily achieved using existing memorization based approaches

e.g. Case-Based Reasoning.
Realization of Proposed Paradigm
Meme Representation

\( x_a \& x_b \): Coordinates of Vertices

\[ M = LL^T \]

\[ d_M(x_a, x_b) = \sqrt{(L^T x_a - L^T x_b)^T(L^T x_a - L^T x_b)} \]
Realization of Proposed Paradigm
Learning Operator

Hilbert-Schmidt Independence Criterion (HSIC)
(Gretton et al., 2005)

Maximize the dependency between $K$ and $Y$

\[
\max_K \quad tr(HKH^T) \\
\text{s.t.} \quad K = X^T \ast M \ast X \\
D_{ij} > D_{iq}, \forall (i, j, q) \in \mathcal{N}, K \succeq 0 \\
H = I - \frac{1}{n}1^T1
\]
Realization of Proposed Paradigm
Selection Operator

\[ \{M_1, M_2, \ldots, M_z\} \]

\[ \mu_i \geq 0, \sum_{i=1}^{z} \mu_i = 1 \]

Selected Meme(s)

\[
\begin{align*}
\mu_1 &= 0.7 \\
\mu_2 &= 0.3 \\
\mu_3 &= \mu_4 = \mu_5 = 0
\end{align*}
\]
Realization of Proposed Paradigm
Selection Operator

HSIC and Maximum Mean Discrepancy (MMD) criteria (Borgwardt et al., 2006)

\[
\begin{aligned}
\max_{\mu} & \quad \text{tr}(HKHY) + \sum_{i=1}^{z} (\mu_i)^2 \text{Sim}_i \\
\text{s.t.} & \quad M_t = \sum_{i=1}^{z} \mu_i M_i, \mu_i \geq 0, \sum_{i=1}^{z} \mu_i = 1 \\
& \quad K = X^T M_t X, K \succeq 0
\end{aligned}
\]

**HSIC: Clustering**

Similarity measure between two given problem instances, defined by MMD and difference in vehicle capacity.

**MMD:**

\[
\| \frac{1}{n_s} \sum_{i=1}^{s} \phi(x_i^s) - \frac{1}{n_t} \sum_{i=1}^{t} \phi(x_i^t) \|
\]

**Sim}_i = -(\beta \ast MMD_i + (1 - \beta) \ast DVC_i)

**DVC}_i Denotes the discrepancy in vehicle capacity between any two problem instances.
Realization of Proposed Paradigm

Variation Operator

\[ \begin{align*}
\mu_1 &= 0.7 \\
\mu_2 &= 0.3 \\
\mu_3 &= \mu_4 = \mu_5 = 0
\end{align*} \]

\[ M_t = \sum_{i=1}^{z} \mu_i M_i \quad \text{Generalization} \]
Realization of Proposed Paradigm

Imitation Operator

From a given New Unseen Problem:

(a) Original task distribution $X_{\text{new}}^{j}$

K-Means on $X_{\text{new}}^{j}$

(b) Knowledge transformed task distribution $X_{\text{new}}^{j}$

(c) K-Means clustering on knowledge transformed task distribution

$X_{\text{new}}^{j} = L^T X_{\text{new}}^{j}$

PDS($X_{\text{new}}^{j}$)

(d) Obtain service orders via pairwise distance sorting (PDS)
Summary on Realization of Proposed Search Paradigm

\[ P_{\text{old}} = \{ p^i \mid i = 1, \ldots, n \} \]

\[ P_{\text{new}} = \{ p^j \mid j = n + 1, \ldots, \infty \} \]

\[ S_{\text{new}}^j = \{ s_g \mid g = 1, \ldots, \text{PopSize} \} \]

Learning via statistical dependence

Knowledge Pool
\[ \text{SoM} = \{ M_i \mid i = 1, \ldots, n \} \]

Selection via \( MMD \)

\[ M_s \]

Variation via \( \text{Generalization} \)

\[ M_r \]

Imitation via \( \text{Clustering and Pairwise Distance Sorting} \)

\[ S_{\text{old}} = \{ s^i \ast \mid i = 1, \ldots, n \} \]

Proposed Approach
Empirical Study

DataSet: Augerat, CE and Christofides CVRP benchmark set
egl CARP benchmark

Evolutionary Solver: CVRP: CAMA, Chen et al. (2012)
CARP: ILMA, Mei et al. (2009).

CAMA and ILMA with Variants of Population Initialization Procedures
- CAMA, ILMA (Random + Heuristics)
- CAMA-R, ILMA-R (Random)
- CAMA-M, ILMA-M (Memes Biased Solutions)
Empirical Study

CVRP & CARP

Search Efficiency

(a) A-n69-k9

(b) B-n41-k6

(a) E3B

(b) S1B
Wilcoxon rank sum test with 95% confidence level has been conducted on the experimental results.
Wilcoxon rank sum test with 95% confidence level has been conducted on the experimental results.
Empirical Study
Insights on a Knowledge Biased CVRP Solution
Memetic Search with Inter-Domain Learning
Realization between CVRP & CARP

An illustration of CVRP and CARP instances and their respective optimized solutions.
Memetic Search with Inter-Domain Learning
Realization between CVRP & CARP

Common knowledge exists in these two problem domains

Pairwise Distance Distributions
Finding a Common Problem Representation

Idea: Customers in CVRP and CARP who bear similar local geometry should be closed to each other in the new common space.

Solution: Derive the mapping from CARP to CVRP and use the latter as the common representation.
Memetic Search with Inter-Domain Learning
Realization between CVRP & CARP

Problem Representation

CVRP instance

CARP instance

Solution Representation

CVRP optimized routes

CARP optimized routes

Knowledge Meme $M_{cvrp}$ and $M_{carp}$

CVRP + $M_{carp}$

CARP + $M_{cvrp}$

High Quality Solution Routes

Memes

Common Representation
Summary on Realization of Proposed Memetic Search with Inter-Domain Learning (CVRP to CARP), (CARP to CVRP)
Empirical Study on CVRP and CARP benchmarks

Some Fundamental Questions:

- Can evolutionary optimization benefit from knowledge meme across problem domains?

- How do different knowledge memes across problem domains influence the evolutionary search?

- What knowledge meme across problem domains would be useful for enhancing evolutionary search?
Can evolutionary optimization benefit from knowledge meme across problem domains?

**Observations**

Evolutionary search can benefit from different but related problem domains.

**Graphs:**
- A-n54-k7
- c199
- CARP
- E3C
- S2C
- CVRP
How do different knowledge memes across problem domains influence the evolutionary search?

OBSERVATIONS

Different knowledge memes introduce unique biases in evolutionary search.
What knowledge meme across problem domains would lead to enhanced evolutionary search?

Diverse discrepancy existed between the respective problem instances.
Conclusions

Towards Nature-Inspired Memetic Computational Search Paradigm

Viewing Memes as Truly Building Blocks that culminate as recurring information patterns & latent structures,

i.e. Higher-Order Representation?
Adaptation, Learning, and Optimization
Founding Series Chief Editor:
Ong Yew-Soon
Lim Meng-Hiot
ISSN: 1867-4534

Proceedings in Adaptation, Learning and Optimization
Founding Series Chief Editor:
Lim Meng-Hiot
Ong Yew-Soon
ISSN: 2363-6084

About this series
The role of adaptation, learning and optimization are becoming increasingly essential and intertwined. The capability of a system to adapt either through modification of its physiological structure or via some revalidation process of internal mechanisms that directly dictate the response or behavior is crucial in many real world applications. Optimization lies at the heart of most machine learning approaches while learning and optimization are two primary means to effect adaptation in various forms.