Humanized computation and interactive evolution

Evelyne Lutton

MIA dpt, UMR GMPA-Grignon, MALICES Team

http://evelyne-lutton.fr
Crowdsourcing

Daren C. Brabham, in 2008: crowdsourcing is an «online, distributed problem-solving and production model.»

- The human brain is used as a computation unit.
- The solution emerges from the crowd.

It is also possible to mix computer-based and human-based computing ...
Open questions

• How to formulate the problem to be solved by the crowd?
  – Splitting the problem into easier subtasks?
  – Emergent global task?
  – Complex systems approach?

• How to manage communication?
  – Competition / cooperation issues?
  – Stigmergy, swarm-like behaviour?
• Humanized computation, interactive evolution.

• A single user in the loop: an example from art.
  – The role of user.
  – The user fatigue issue.

• « man-hills »: an example from e-leaning.
  – Ant colony principle.
  – Adaptation using a chess Elo notation.
Hideyuki Takagi: «Combination of human factor with computational intelligence. Directly embed a human as a black box into a target system. »

- Human capabilities combined with autonomous computations for dealing with complex problems.

- Optimisation / Interaction / Visualisation
Combination of expert interaction and autonomous computations

Heterogeneous data and knowledge

Inverse problems, multi-objective optimisations.

Evolutionary Computation is a convenient framework

Issues

* Efficient visualisation for complex multidimensional data,
* User fatigue.

Optimisation / Interaction / Visualisation

Complex systems

Selection

Mutation

Crossover

Elitism

Visualisation

PARENTS

OFFSPRING

Extraction of the solution

Initialisation

Optimisation

Interac/on
**Darwinism, Evolutionnism**

- Variations, macroscopic and microscopic, inside species,
- Struggle for life,
- Natural selection: expansion of the lineage that has a useful variation in his environment.

**Evolutionary algorithms**

Global optimisation technique:
Mimics the collective learning ability (adaptation) of natural populations.
Evolutionary Optimisation

Population N

- Parents
- Selection of two parents
- Crossover
- Mutation of the two children

Population N+1

Initial Population

Evolved Population

One individual = one solution
The human in the loop ...

Additional constraints:
- slows down the computation rythm,
- interaction and visualisation issues
User interaction issues

**FEW INTERACTION, REPETITIVE TASKS:** Tiredness, Boredom.

**TOO MUCH CHOICES:** Puzzlement in an Ocean of Information

**trade-off:**

- Fluidity / adequation of visualisations
- Navigation freedom
- Focus on interesting areas

(interaction) (exploration) (optimisation)
Mono- and Multi-users interactions

- Synchronizing users interactions and computations.
- Fluid visualisation and interaction.

User fatigue = “user bottleneck” of IEC

**Proposed solutions:**
- Reduce the number of questions (population size + nb of generations).
- Choose a model carefully (visit only interesting areas of the search space).
- Help the user (automatic learning, surrogate models).
Fractals are mathematical models of some natural shapes
Fractals are obtained as a limit of an iterative process

Sierpinski triangle
The Von Koch snowflake

With 4 elementary transformations only!
The Barnsley Fern

- With 4 linear functions also 😊
Iterated functions systems attractors

With complex functions
Non linear IFS ...

How to explore this space of shapes?

\[ w_1(x, y) = \left( \frac{1 \cdot e^{-1 \cdot \text{atan}(x, y)^2}}{\sqrt{(x - 0.5)^2 + (y - 0.5)^2}} \right) \]

\[ w_2(r, \theta) = \left( 0.5 \cdot (\text{tanh}(0.95 \cdot (0.5 \cdot \text{tanh}(\theta) + 0.5)^{r+\theta}) + 0.5) \cdot r \cos((-0.25 \cdot \cos(-1.18 \cdot \text{atan}(\theta \cdot r)) + 0.75) \cdot (\theta - 2.04)) + 2.04 \right) \]
Random Mutation
Random Crossover
Interactive Mutation
The designer guides the evolution of his population of shapes
Video artwork by Anabela Costa
Interactive Design

• Even very simple evolution schemes allow to obtain results.
• Trade-off between randomness and user-guided search = augmented creativity.
• Lateral thinking and surprise effects.
• The user learns how to use the interactive system, and creates his own way to use the system.
• **What seems important:**
  – Convergence issues have to be revisisted: creation is NOT an optimisation.
  – Exploration components are very important.
  – Giving more and various access to the process reduces “user fatigue”!
A multi-users example: e-learning

- French leading e-Learning company
- More than 7,000 provided items
  - Free navigation
  - "Static" guided navigation

Main goals of the PhD:
- Optimisation of the learning.
- Adaptation of the navigation to the student
- Emergence of new pedagogical paths.
Emergence of paths & adaptability
An ant colony scheme for Paraschool

- Items (exercises, courses) are organised as a graph
- Deposit and evaporation of informations
- Free navigation of students → emergence of new paths
- A huge number of users : 250.000 pupils
Pheromones ($\phi^+ / \phi^-$)

- Allow collective exchange of information
- Two types of pheromones: success / failure
  - Deposit on edges ($\alpha$)
  - Backpropagation ($1/k$)

Emergence of preferred paths
**Pedagogical information**

**Teachers weight (W)**

Representation of pedagogy

Relative weights

Edges creation is possible ...
Principle
Local Fitness

Edges are evaluated using a local fitness function

Learning ratio (A)
   Aim = 60% of success

Preferred path by students (P)
   Influence of edges pheromones

Confrontation with teachers (C)
   Teacher rating, for favouring some paths

\[
\text{Fitness} = w_1 C + w_2 P + w_3 \phi^+ + w_4 \phi^-
\]

Selection based on evolutionary search schemes
A community of connected students is not an ant colony

1. Irregular activity (holidays)  
   \[ \rightarrow \text{stigmergic data get lost} \]
2. Heterogeneity of agents behaviour  
   \[ \rightarrow \text{less control} \]
3. No altruism  
   \[ \rightarrow \text{exploration is less efficient} \]
4. There is no clear collective objective  
   \[ \rightarrow \text{selfish behaviour} \]

New Paradigm = "Man-Hill"

- Erosion
- Individual features
• Evaporation according to time
  
  Low activity → loss of information
  
  High activity → overestimation of $\phi$

• Evaporation according to activity = erosion

  Each time an item is tested

  Success/ Failure → $\phi^+ / \phi^-$ are updated
  $\Phi$ is decreased
Goal → Adapt the pedagogical path to each student

Additional parameters:

- **History parameter**
  to avoid exercises already done
  → Evaporation process

- **Agenda parameter**
  to propose exercises selected by the teacher

- **Level parameter**
  to propose level-adapted exercises
  but there is no rating system on Paraschool ...
Teacher's Rating

Pourcentage d'échec des élèves
Competition of Student (Ss) against Item (Si)

- Expected result: \( R_e = \frac{1}{1 + \exp\left(\frac{S_i - S_s}{400}\right)} \)

- Real result: \( R_r = \{0, 0.5, 1\} \)

- Updated ELOs: \( S_s = S_s + K \cdot (R_r - R_e) \)
ELO-like evaluation of items and students

For students
  Measurement of level and evolution

For the pedagogical team
  Automatic rating of items
  Feed-back
  Detection of semantic and pedagogical errors

Man-Hill approach
  Allow proposing items adapted to level of students.
1. Sub pools
   - Different classes.
     + 95% of the students stay in their classes.
     + Elo rating not updated if $\text{Class}_{\text{item}} \neq \text{Class}_{\text{student}}$ (5%)
   - Two particular sub pools
     + Items' Sub pool / Students' Sub pool

2. Turnover
   - Students entering/leaving the system
   - Periodicity: school year
Reducing deflation of items' rating

- Reset the rating of items each year. 😞
- Stop evolution before deflation (Freeze)

- Use a different rating scheme for items (PERI)
New rating system using the structure of Paraschool

- **Sedentary** group = items → needs a more stable rating
- **Nomad** group = students → dynamic **Elo rating**

**Probability based Elo Rating for Items: PERI**
- Elo rating → Too adaptive → Deflation
- Freezing method → Too static

**Item success rate**
- Elo Theory: The difference of Elo gives the probability of success.
- PERI: success rate of students on an item allow to update the rating:

\[
S_{item} = S_{ref} + 400 \log_{10} \left( \frac{1 - \%\text{success}}{\%\text{success}} \right)
\]
• PERI & Freezing: More stability on Items' ratings
• PERI: Smaller oscillations on Students' ratings
CONCLUSION

• The translation from single to many users is not a simple problem of scale ...
• Emergent effects appear, difficult to understand and control.  
  A crowd-based system is a complex system. 😊

• Homogeneity is a big issue : example of the « man-hill » with two subpopulations.
• Psychological attitude is important : example of user-fatigue.

• Mixing user interactions and autonomous computations is another huge challenge ...
  Machines can help humans ... and reversely !
Thank You!