

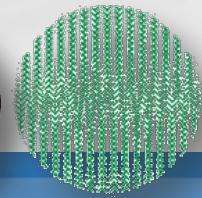
Promoting Diversity in Evolutionary Optimization: Why and How

Giovanni Squillero

giovanni.squillero@polito.it

Alberto Tonda

alberto.tonda@grignon.inra.fr



INRA

Saturday, September 17th 2016



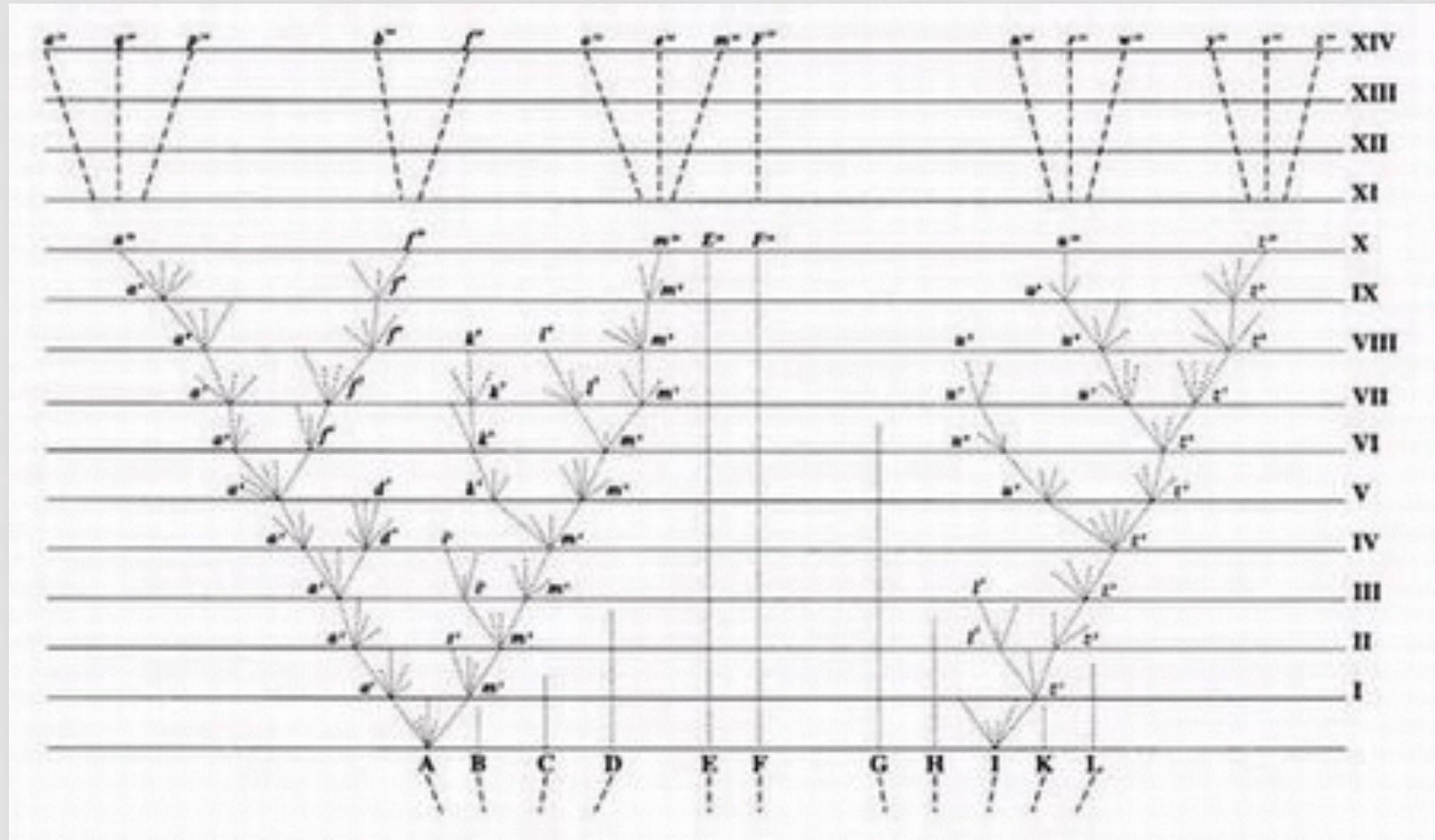
Outline

- Generic EA
- Divergence of character in natural and artificial evolution
- Background (diversity and similarity, ...)
- Mechanisms for promoting diversity
- Hints and tips
- Conclusion

Real world (Galapagos)

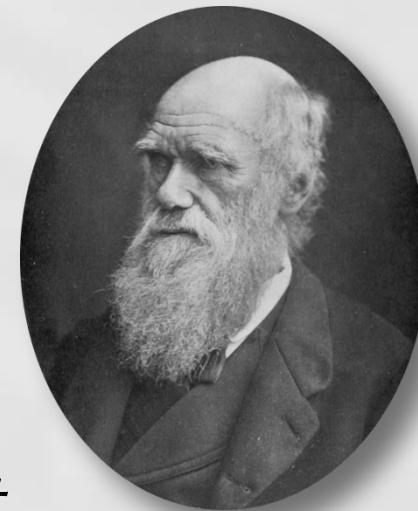


Darwin's tree of life



The only illustration in *On the Origin of Species* by Natural Selection (1859)

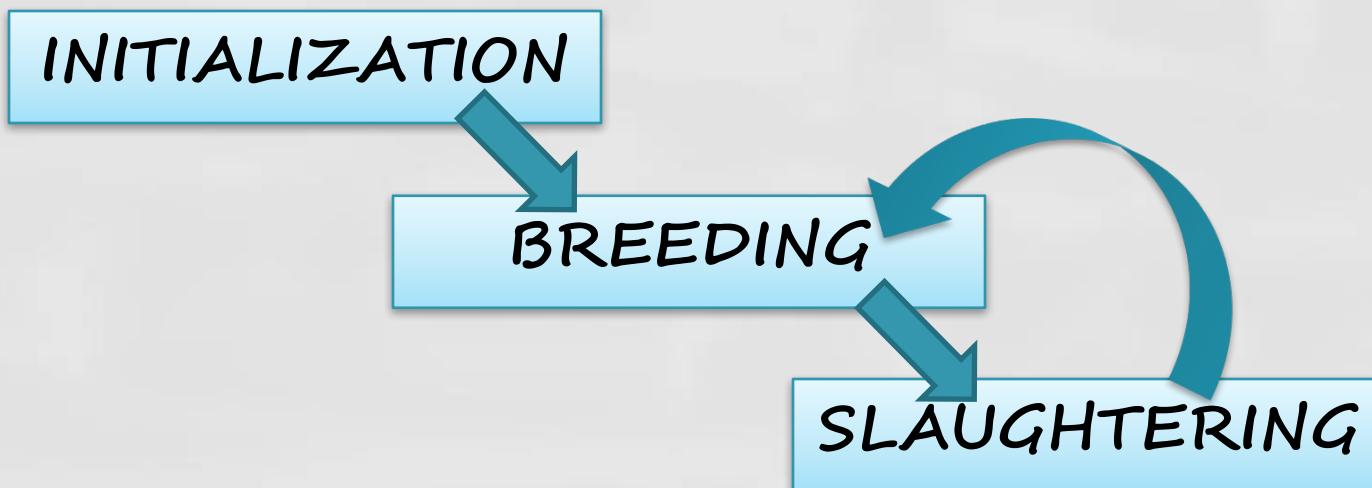
Divergence of character



- “Great diversity of forms in nature”
- “The principle, which I have designated by this term, is of high importance, and explains, as I believe, several important facts”
 - “The principle of divergence causes differences, at first barely appreciable, to steadily to increase, and the breeds to diverge in character, both from each other and from their common parent”
 - “The varying descendants of each species try to occupy as many and as different places as possible in the economy of nature”

Evolutionary computation

- A rough idea about “what” an evolutionary algorithm is



- Note: Optimization, not artificial life!

Evolutionary algorithms



Premature convergence

- I.e., the tendency of an algorithm to converge towards a point where it was not supposed to converge to in the first place
- Probably an oxymoron
- Holland's “Lack of speciation”
- EAs general inability to exploit environmental niches

divergence of character

vs.

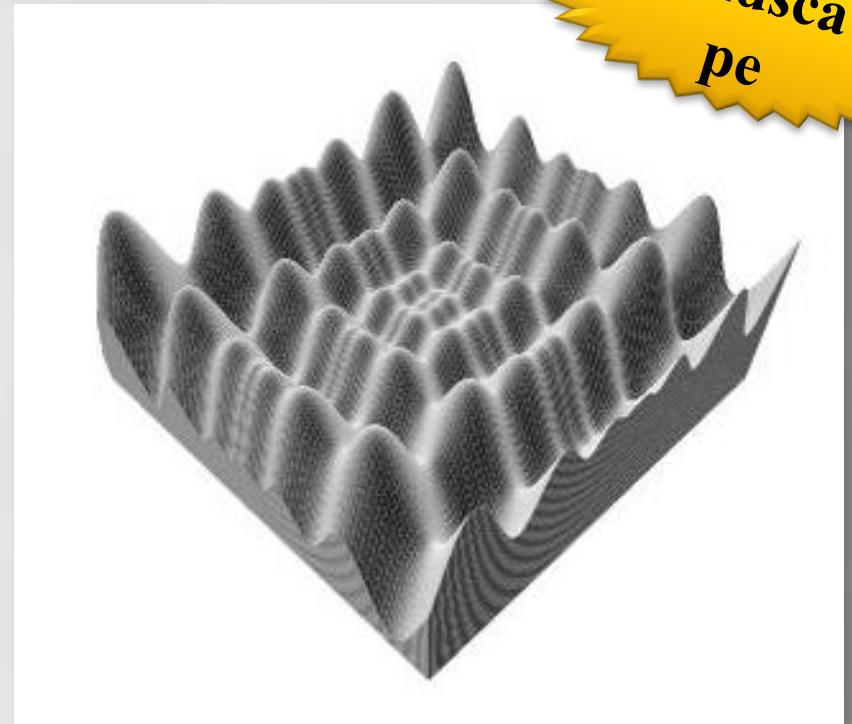
premature convergence

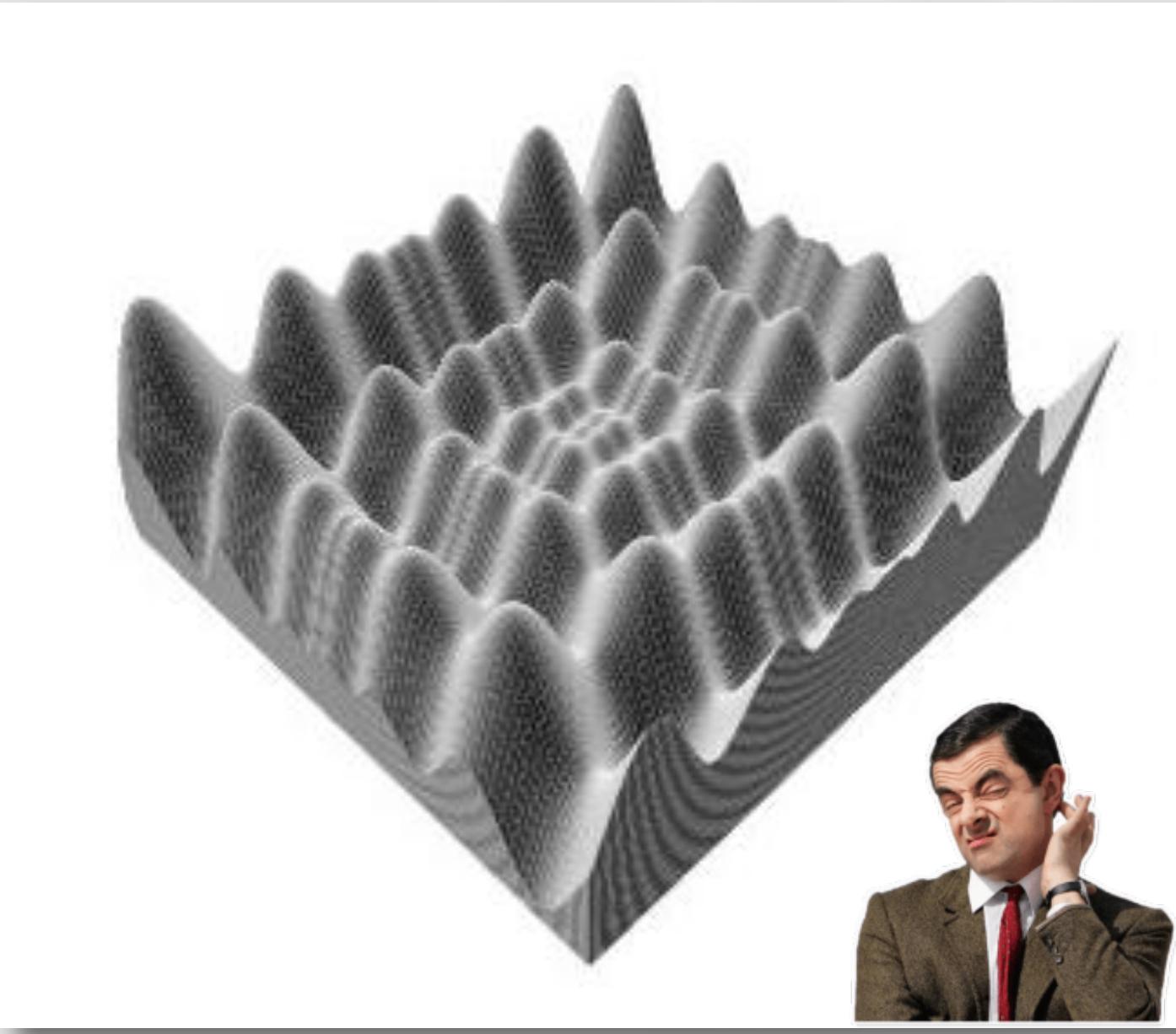
Divergence of character

- “The basic point of the principle of divergence is **simplicity itself**: the more the coinhabitants of an area differ from each other in their ecological requirements, the less they will compete with each other; therefore natural selection will tend to favor any variation toward greater divergence.”



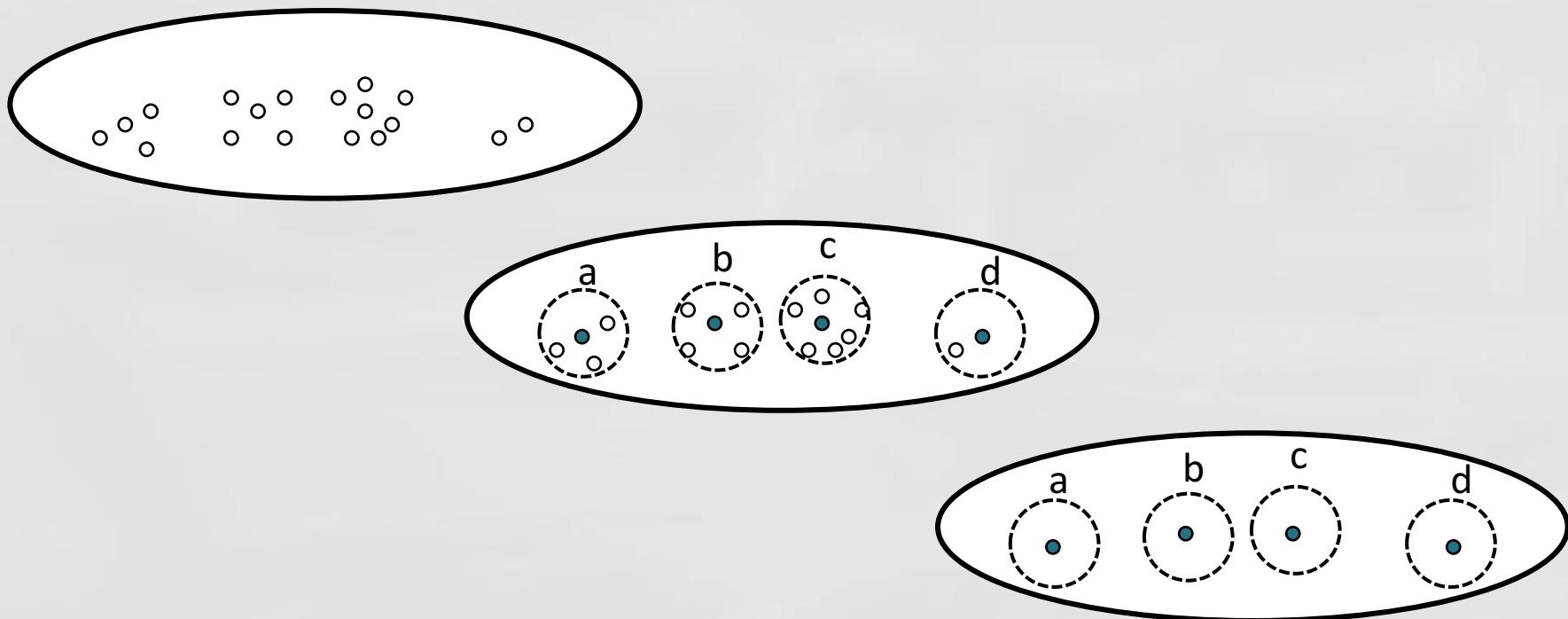
Environment vs. Fitness function





Niches

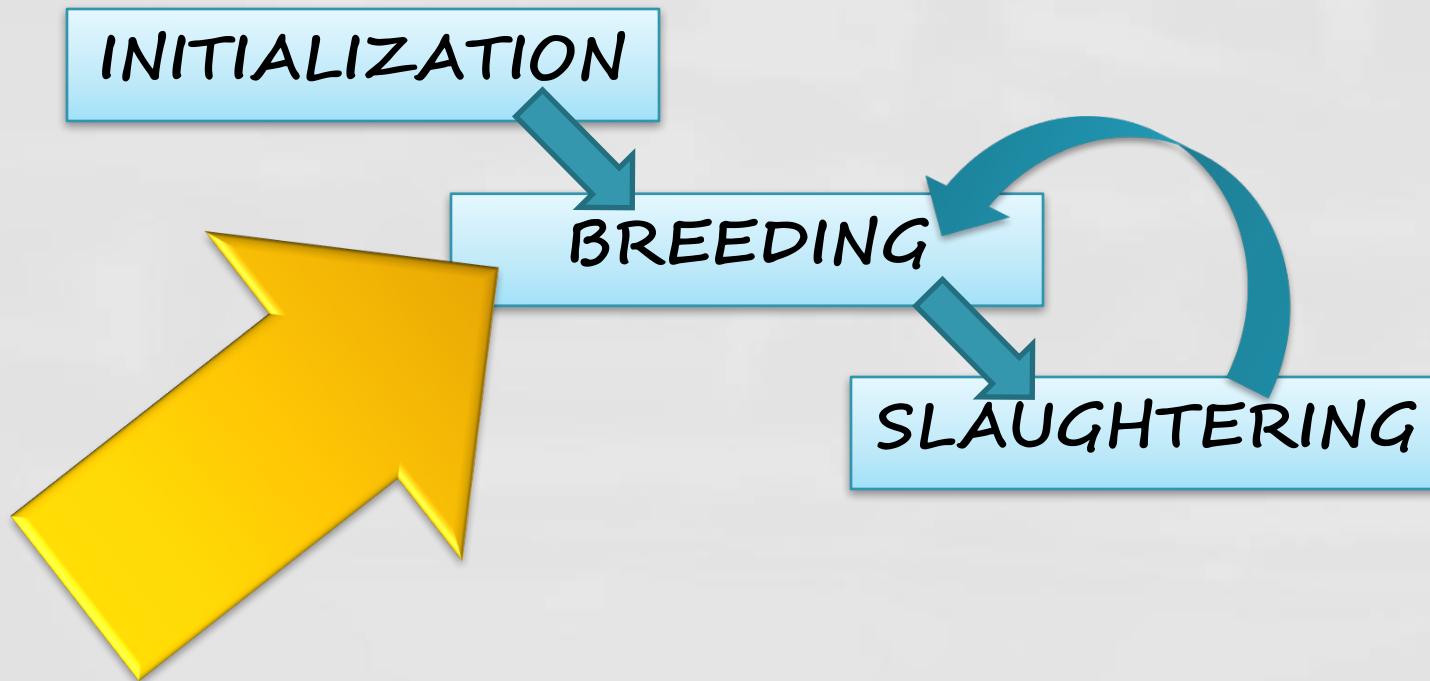
- Niche: subspace in the environment with a finite amount of physical resources that can support different types of life



Niches

- Niches favor the divergence of character
- Niches and speciation
- How to create “niches” in EAs since the environment is missing?

A Very Generic EA



Exploration vs. Exploitation

- Recombination
 - mixes together two or more solutions to create the offspring
 - associated with the idea of *exploration*
- Mutation
 - performs a (usually small) change in an individual
 - associated with the idea of *exploitation*



Exploration vs. Exploitation

- When all parents are very similar, the effectiveness of recombination is limited
- The ability to explore remote parts of the search space is impaired
- “Conventional wisdom suggests that increasing diversity should be generally beneficial”



Exploration vs. Exploitation

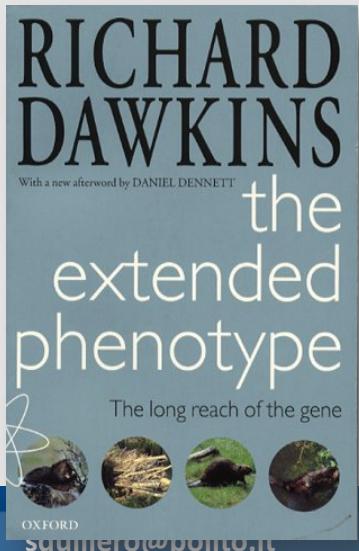
- When all parents are **very similar**, the effectiveness of recombination is limited
- what is the definition of “similar”?**
- space is impaired
- “Conventional wisdom suggests that increasing **diversity** should be generally beneficial”

and the definition of “diversity”?



Levels in biology

- **Genotype:** the genetic constitution of an organism
- **Phenotype:** the composite of the organism's observable characteristics or traits
- **Fitness:** individual's ability to propagate its genes (well, almost)

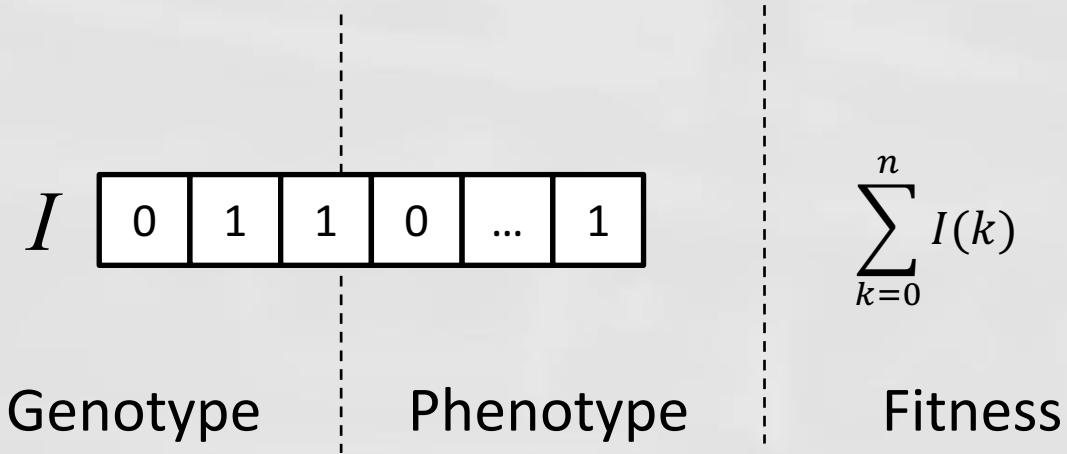


Richard Dawkins
The Extended Phenotype: The Long Reach of the Gene
Oxford University Press, 1982 (revised ed. 1999)

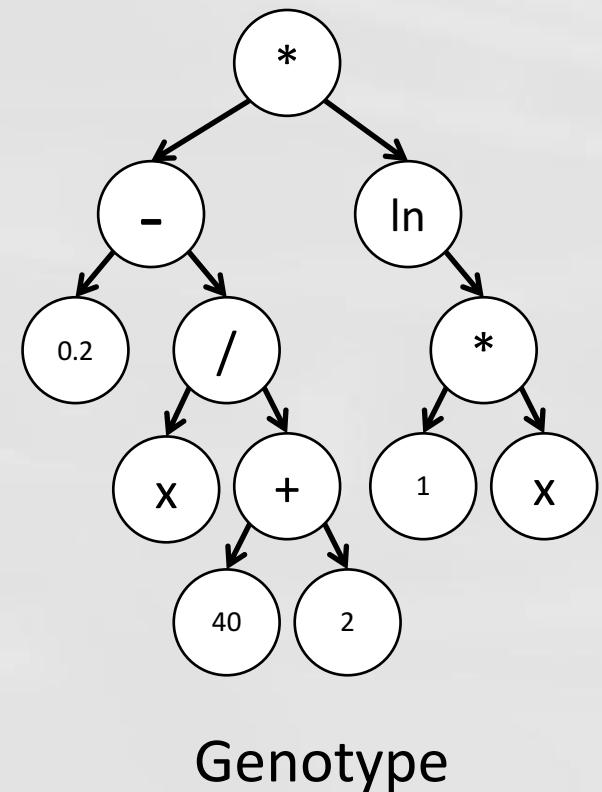
Levels in EC (a modest proposal)

- **Fitness:** how well the candidate solution is able to solve the target problem
- **Genotype:** the internal representation of the individual, i.e., what is directly manipulated by genetic operators
- **Phenotype:** the candidate solution that is encoded in the genotype
 - the intermediate form in which the genotype needs to be transformed into for evaluating fitness
 - if genotype can be directly evaluated: genotype and phenotype coincide

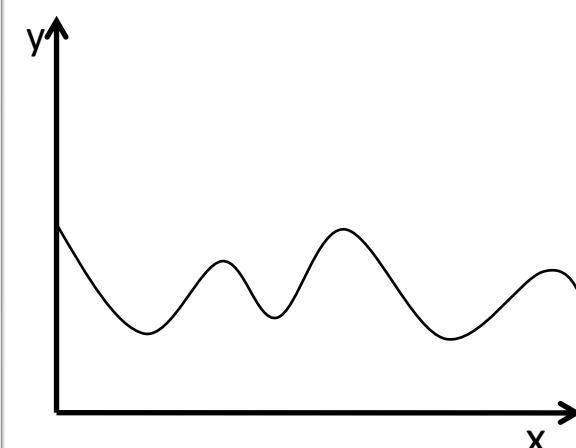
Levels in EC (GA)



Levels in EC (GP & LGP)

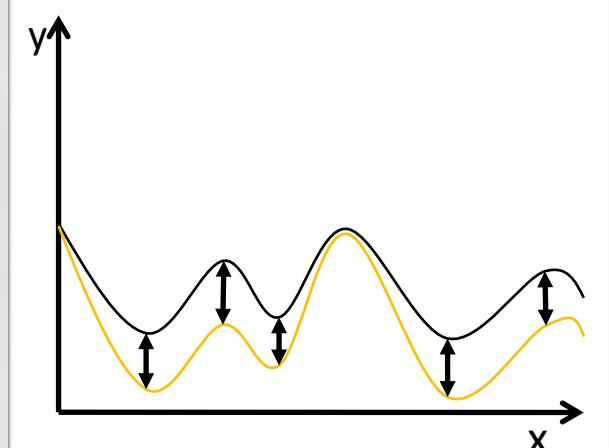


$$f(x) = \left(0.2 - \frac{x}{42}\right) \cdot \ln(x)$$



Phenotype

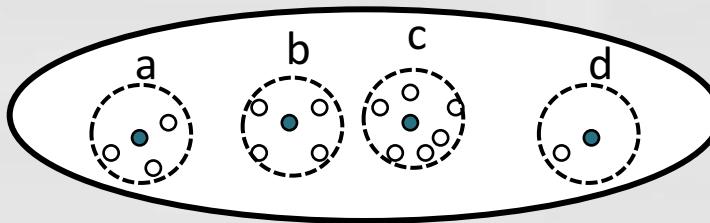
$$\text{Fitness} \approx \int_{x=0}^E |f(x) - g(x)|$$



Fitness

Niches in EA

- Niching: grouping similar individual
 - similar spatial positions (i.e., islands)
 - similar genotypes (i.e., niching)
 - similar phenotypes
- Several approaches are based on niching



Detecting clones

- Detecting whether two individuals are clones, i.e., identical, is often an easy task at any level



Measuring diversity

- Diversity \Rightarrow distance metric: *how far* the individual is
 - from (a subset of) the whole population
 - from a single individual
- Diversity \Rightarrow property of the population
- But, at what level?
 - Phenotype
 - Genotype
 - Fitness



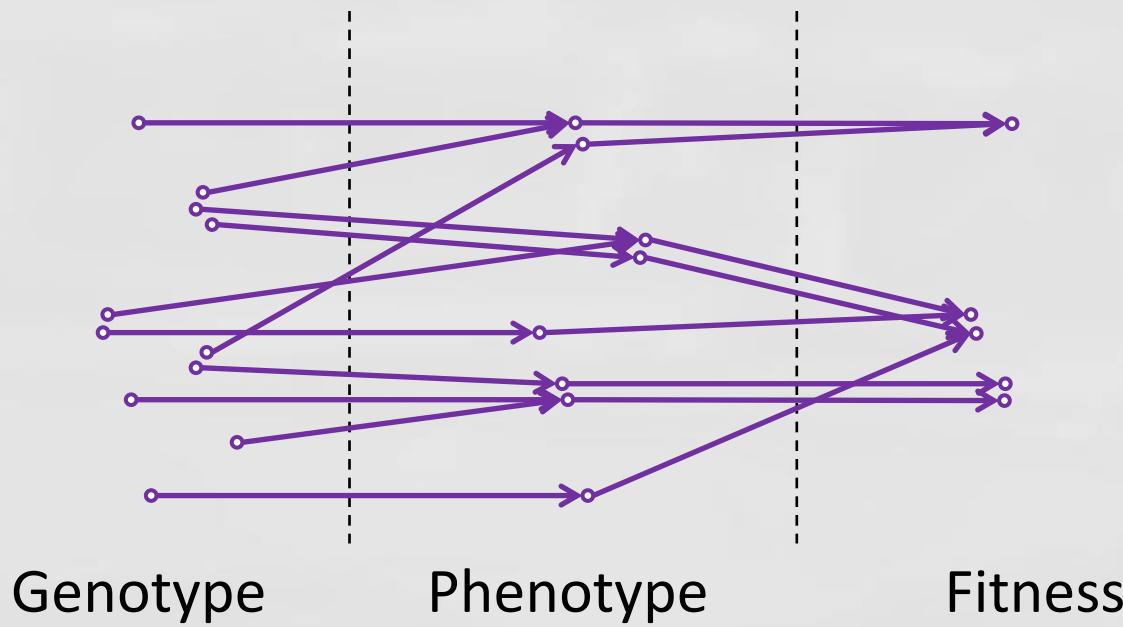
Measuring diversity

- Different fitness values imply different phenotypes, different phenotypes imply different genotypes

$$F_x \neq F_y \Rightarrow P_x \neq P_y \Rightarrow G_x \neq G_y$$

Measuring diversity

- What about “diversity”?
- Locality principle
- Rechenberg’s *strong causality*



Measuring diversity

- Fitness
 - Usually trivial
- Phenotype
 - Usually ad-hoc
- Genotype
 - Different genotypes in the population
 - GP subtree frequency
 - Edit distance (a.k.a., Levenshtein distance)
 - Entropy and free energy

Outline

- Generic EA
- Divergence of characters evolution
- Background (diversity and similarity, ...)
- Mechanisms for promoting diversity
- Hints and tips
- Conclusion

*What has been
proposed
to alleviate it?*

End goal vs. Means goal

- The **end goal** in optimization is reaching better solutions in less time
- Promoting diversity has often been seen as the key factor to improve performances
- Promoting diversity is a mere **means goal** (yet a quite important one)
- No distinction is made here whether the means goal is
 - preserve existing diversity
 - increase diversity

How diversity is promoted (practice)

- Fitness scaling
- Fitness holes
- Tweaking selection mechanism
- Adding selection mechanism
- Multiple populations
- Population topologies
- ...

In theory there is no difference between theory and practice



How diversity is promoted (theory)

- A methodology for promoting diversity alters the selection probability of individuals

$$\bar{p}_{x|\Psi} = p_{x|\Psi} \cdot \xi(x, \Psi)$$

- **Mere definition:** we do not imply that a mechanism operates *explicitly* on the selection operators
- **But** the *effects* on selection probabilities are assessed to classify it

How diversity is promoted (theory)

individual

set of individuals
(may be empty)

$$\bar{p}_{x|\Psi} = p_x \cdot \xi(x, \Psi)$$

selection probability of individual x given that all individuals in set Ψ are also chosen

selection probability of individual x

corrective factor

Relevant characteristic

- Lineage (**LIN**)
- Phenotype (**PHE**)
- Genotype (**GEN**)
- ~~Fitness~~ (used as a proxy for either phenotype or genotype)

$$\xi(x, \Psi)$$

Lineage-based methodologies

- The value of $\xi(\circ)$ does not depend on individual structure nor behavior, but it can be determined considering circumstances of its birth (e.g., time, position)
- LBMs can be applied to any kind of problem, even in addition to other diversity preservation methods

Genotype-based methodologies

- Particularly effective when it is possible to define a sensible distance between genotypes
- Often used to
 - avoid overexploitation of peaks in the fitness landscape
 - promote the generation of new solutions very far from the most successful ones
 - preserve variability in the gene pool

Phenotype-based methodologies

- Usually impractical
- Sometimes fitness distance can be used as a proxy for phenotype distance (multi objective EAs, or many objective EAs)

Type of selection

- Parent selection (α or ϵ)
 - Usually non-deterministic
- Survival selection (ω or η)
 - Usually deterministic

$$\bar{p}_x | \Psi$$

Diversity Promotion in EA



ELSEVIER

Information Sciences

Volume 329, 1 February 2016, Pages 782–799

Special issue on Discovery Science



Divergence of character and premature convergence: A survey of methodologies for promoting diversity in evolutionary optimization [☆]

Giovanni Squillero  ^a,  , Alberto Tonda^b, 

DOI: [10.1016/j.ins.2015.09.056](https://doi.org/10.1016/j.ins.2015.09.056)



Island model

- Recipe [LIN $\alpha\omega$]

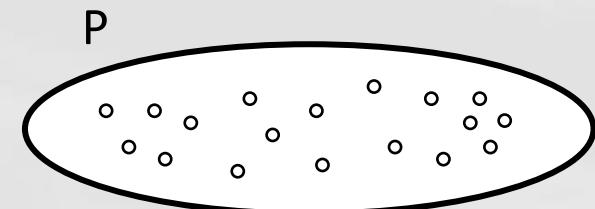
- The population is partitioned into sub-populations
 - Only local interactions are allowed
 - Periodically, individuals are moved between sub-populations (**migrants**)

- Rationale

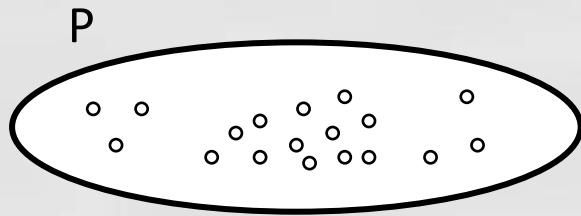
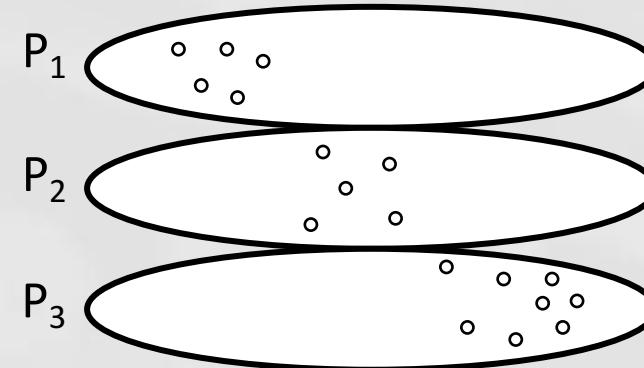
- Different populations may explore different parts of the search space
 - ... but global interactions can be useful



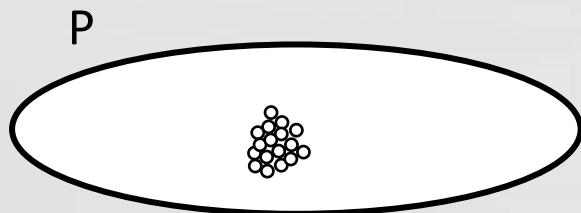
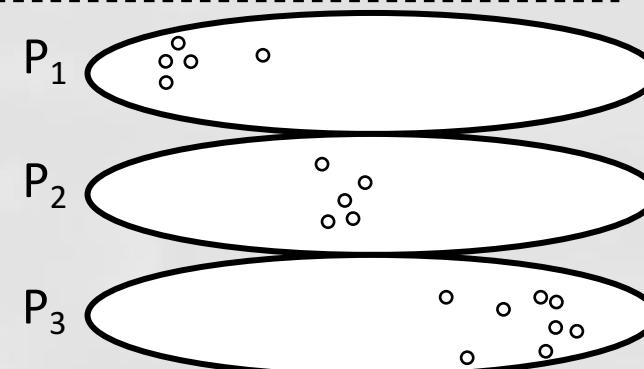
Island model



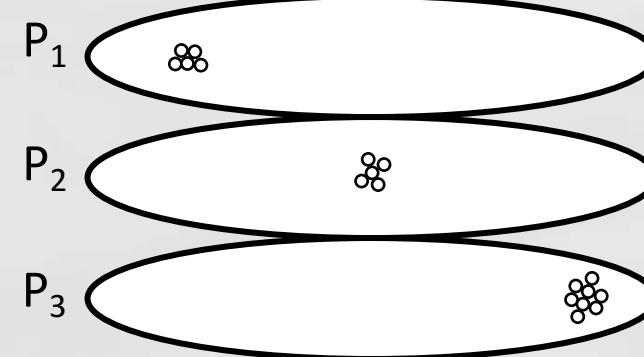
$T = t_0$



$T = t_k$



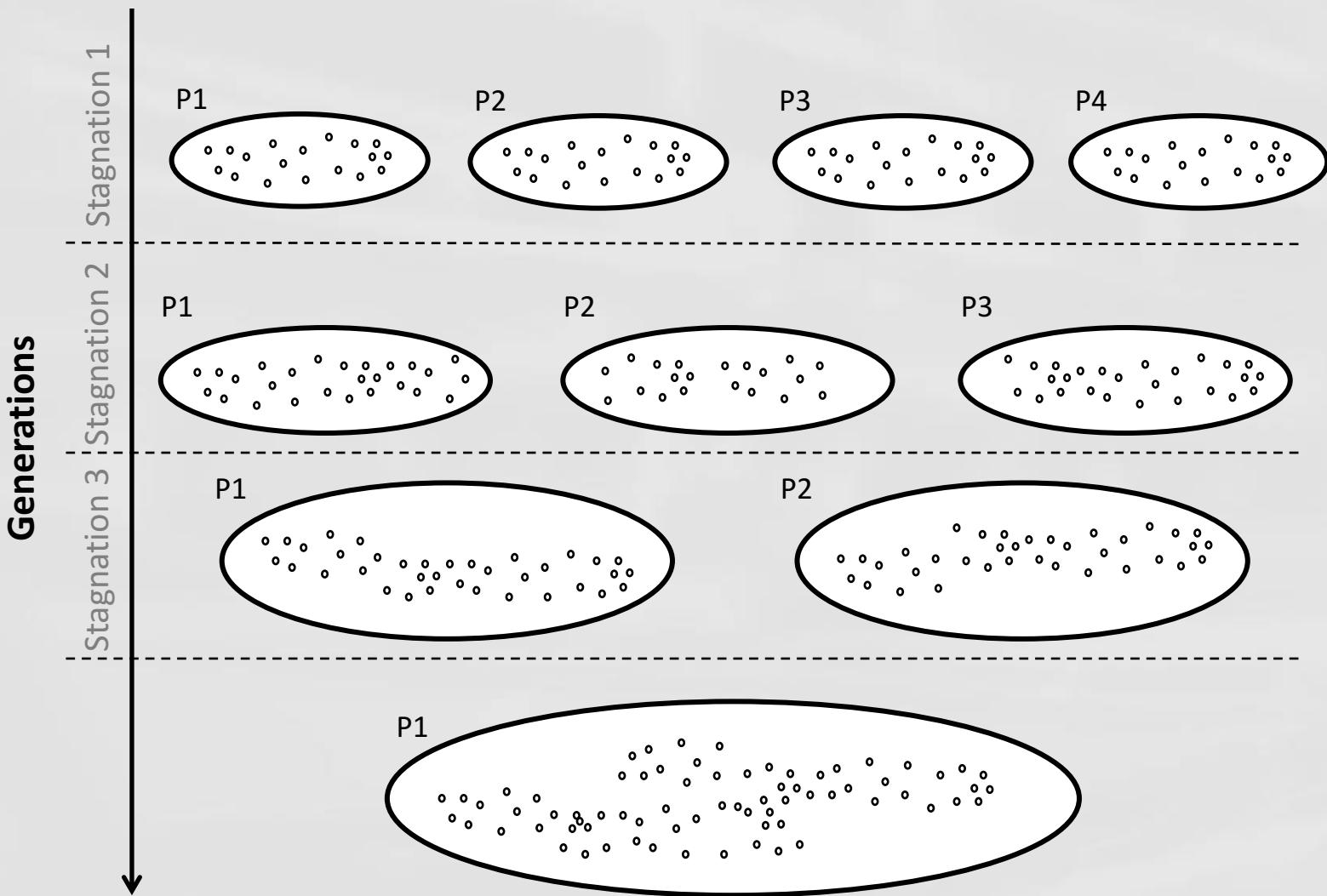
$T = t_N$



Segregation

- Recipe [LIN $\alpha\omega$]
 - The population is partitioned into N sub-populations
 - Only local interactions are allowed
 - Upon stagnation, the N sub-populations are merged into $N-1$ sub-populations
- Rationale
 - Same as island models
 - The selective pressure decreases during evolution

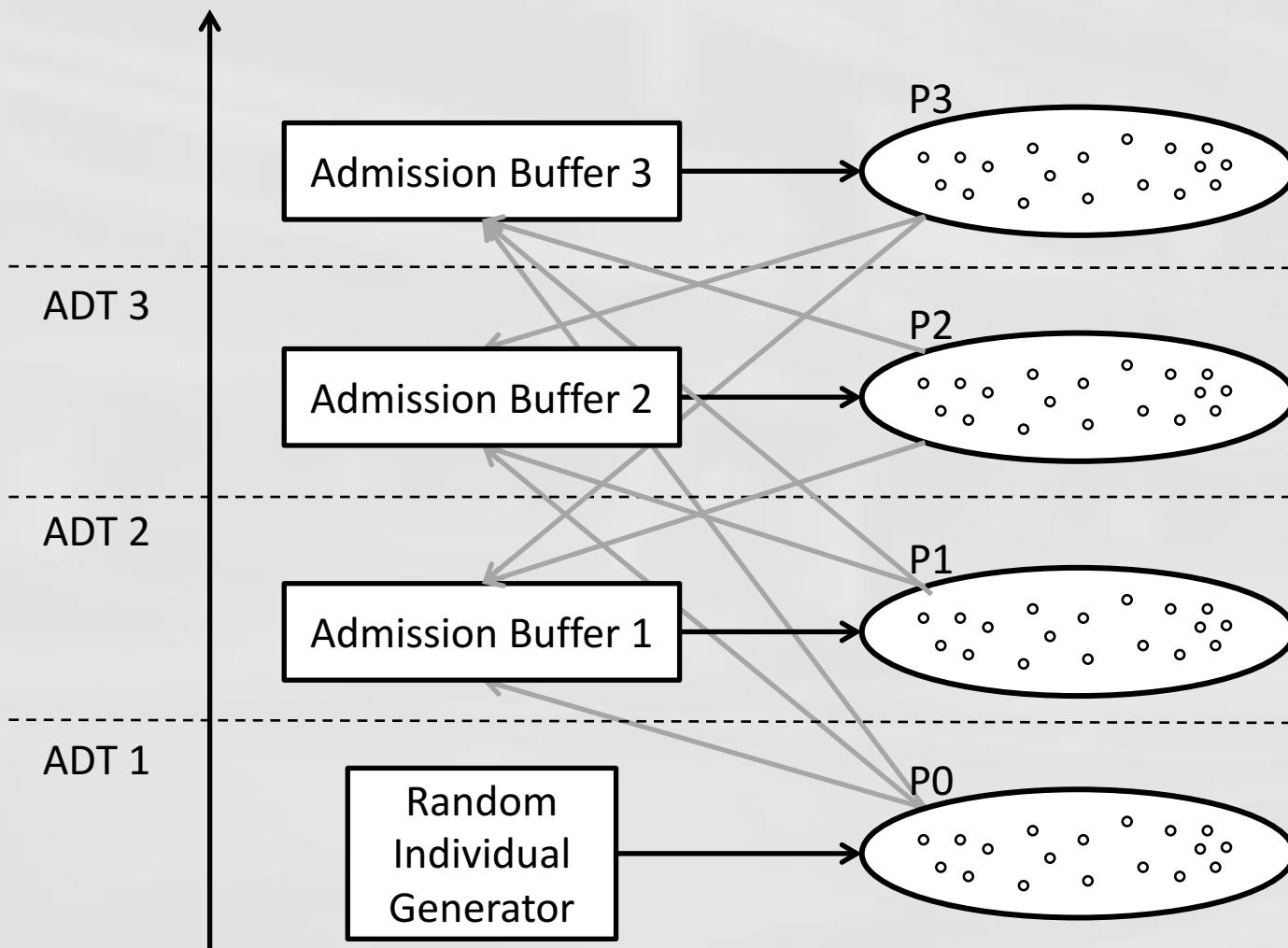
Segregation



Hierarchical fair competition

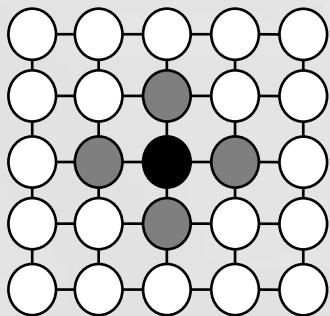
- Recipe [PHE $\alpha\omega$]
 - The population is partitioned into sub-populations with similar fitness
 - Only local interactions are allowed
 - The offspring is promoted or demoted according to fitness
 - New random individuals are constantly generated
- Rationale
 - Hard niching with implicit neighborhood
 - Reduce competition between newborns and already optimized individuals (ladder)

Hierarchical fair competition

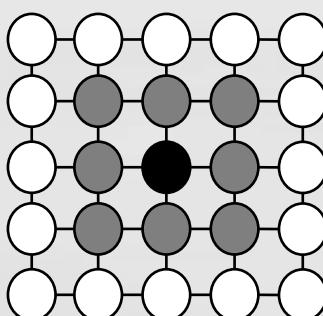


Cellular EA

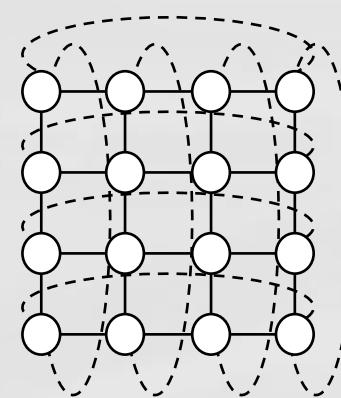
- Recipe [LIN $\alpha\omega$]
 - Fixed topology (lattice)
 - Only interactions between neighbors are allowed
- Rationale
 - Limiting interaction could defer the takeover of the population by clones of the fittest individual



Linear-5 (L5)



Compact-9 (C9)



Deterministic crowding

- Recipe [LIN et al]
 - Offspring compete against parents for survival
- Rationale
 - Flexible niching with implicit neighborhood
 - Parents and offspring occupy the same niche
 - No need for evaluating the similarity

Allopatric selection

- Recipe [LIN ~~etw~~]
 - The whole offspring compete for survival
- Rationale
 - Flexible niching with implicit neighborhood
 - No need for evaluating the similarity
 - Genetic operators that create large offspring can be exploited without the risk for the offspring to invade the population

Fitness Sharing

- Recipe [GEN aw]

- Scale down individual fitness

$$\bar{f}(I_k) = \frac{f(I_k)}{\sum_i sh(I_k, I_i)}$$

- with $sh(x, y)$ depending on the distance between the individuals, and is 0 beyond a fixed radius

- Rationale

- Flexible niching with explicit neighborhood
 - Reduce attractiveness of densely populated area

Clearing

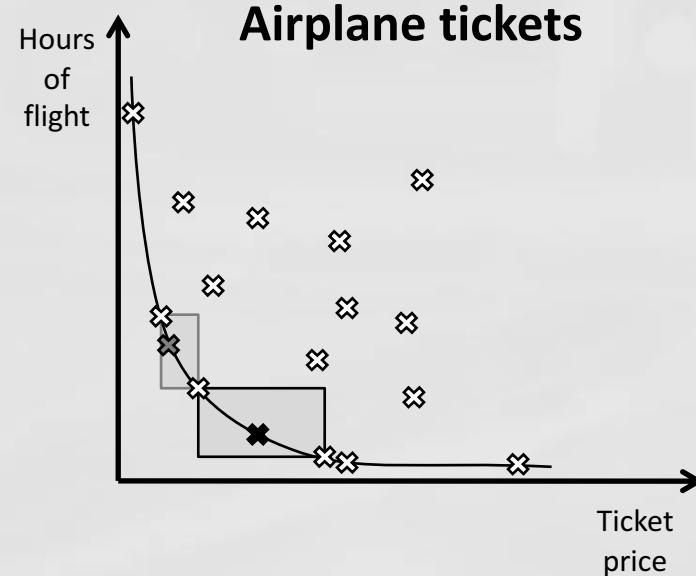
- Recipe [GEN $\alpha\omega$]
 - Inside niches of a certain radius, the best k individuals retain their fitness while the rest are zeroed
- Rationale
 - Flexible niching with explicit neighborhood
 - Set a hard limit to population density

Standard crowding

- Recipe [GEN ~~etw~~]
 - New individuals replace the most similar individual in a random niche of size CF
- Rationale
 - Flexible niching with explicit neighborhood
 - Favor novelty (generational approach)

Crowded-comparison operator

- Recipe [PHE ~~etc~~]
 - Estimate the *free territory* around solutions and favor solutions less crowded regions
- Rationale
 - Smart implementation of artificial niches
 - Requires a strong correlation between phenotype and fitness
 - NSGA-III introduces ϵ -domination (adaptive discretization)



Reference points partitioning

- Recipe [GEN ]
- Population is partitioned using in clusters centered around a set of reference points
- Reference points are initially chosen by the user, then can be dynamically updated
- New individuals compete for survival inside their own niche
- Rationale
 - Flexible niching with explicit neighborhood

Vector evaluated genetic algorithm

- Recipe [PHE ~~etc~~]
 - Divide the mating pool in N parts, each one filled with individual selected on their i -th component of the fitness
 - Alternative: select on a weighted sum, but use different weight sets for the different parts
- Rationale
 - Increase the push towards specialization
- Caveats
 - Only applicable to MOEAs, or when using an aggregate fitness

Lexicase selection

- Recipe [PHE ~~etc~~]
 - Before selection, re-arrange the components of the fitness
 - Compare individual fitnesses lexicographically
- Rationale
 - Increase the push towards specialization
- Caveats
 - Only applicable when using an aggregate fitness

Restricted tournament selection

- Recipe [GEN ω]
 - New individuals compete with the most similar individual in a random niche of size CF
- Rationale
 - Flexible niching with explicit neighborhood

Sequential niching

- Recipe [GEN $\alpha\omega$]
 - The most promising points in the search space after each run are altered so to become less interesting in further executions
- Rationale
 - Avoid over exploitation

Gender

- Recipe [LIN/GEN ~~atx~~]
 - Add gender to individual and enforce sexual reproduction
 - More than two sexes are possible, with different mutation probabilities
 - Gender might be part of the genome or not
- Rationale
 - Prevent crossover between clones
 - Limit interactions between related individuals

Tarpeian method

- Recipe [PHE ω]
 - Randomly kill individual who don't adhere to given standards
- Rationale
 - Note: originally used to prevent bloat
 - Creating dynamic and non-deterministic *fitness holes* may have several beneficial effects, including to promote diversity

Diversifiers

- Recipe [GEN $\alpha\omega$]
 - Detect less populated areas in the search space and try to generate random inhabitants
- Rationale
 - Increase variability in the gene pool regardless the fitness
 - Require a reliable distance metric

Random immigrants

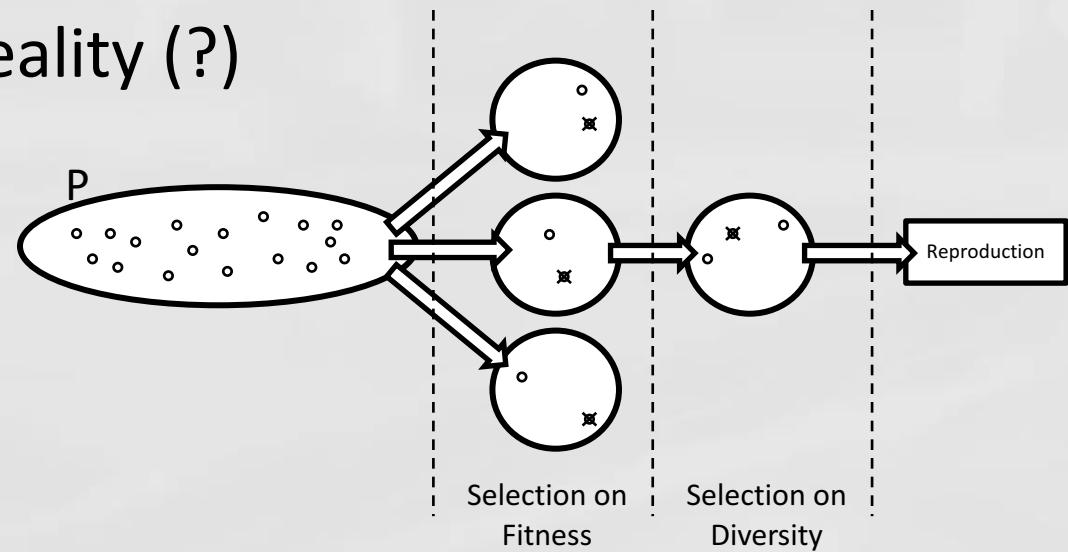
- Recipe [PHE $\alpha\omega$]
 - Periodically insert random individuals in the population
- Rationale
 - Try to introduce novelty
- Caveats
 - Newborns may need to be artificially kept alive when competing against already optimized individuals

Extinction

- Recipe [PHE ]
- Upon convergence (or periodically) remove a significant part of the population
- Then fill up the population with the offspring of the survivors and/or random individuals
- Rationale
 - A gust of fresh air: already optimized individuals are not enough to occupy the whole population and newborns may start exploring new regions
- Caveat
 - Fitness variability used as phenotype variability

Two-level diversity selection

- Recipe [GEN ~~atx~~]
 - Select three individuals using fitness, then pick the two with maximum distance for reproduction
- Rationale
 - Exploit a reliable distance metric to increase the efficacy of crossover
 - Not so far from reality (?)



GDEM – Genetic Diversity Evaluation Method

- Recipe [GEN $\alpha\omega$]
 - Add diversity as an explicit goal and go MO
- Rationale
 - Modify the domination criteria
 - Need a reliable diversity metric
- Historical note
 - See: *Find Only and Complete Undominated Sets* (FOCUS)

Delta entropy and pseudo entropy

- Recipe [GEN ~~and~~]
 - With a certain probability select individuals on their ability to increase the global entropy of the population instead of fitness
- Rationale
 - Not-so-fit individual with peculiar traits should be preserved
 - Measuring the entropy of the population is easier than defining a distance function

Outline

- Generic EA
- Divergence of character in natural and artificial evolution
- Background (diversity and similarity, ...)
- Mechanisms for promoting diversity
- Hints and tips
- Conclusion

Hints and Tips

- Do you really need to promote diversity?
 - Several problems in EA are caused to ill-designed fitness functions
 - Check whether the locality principle holds true
 - Check what happen with multistart



Hints and Tips

- ➊ Do you really need to promote diversity?
- ➋ Use extinction (20m)
 - ➌ Simple n' easy



Hints and Tips

- Do you really need to promote diversity?
- Use extinction (20m)
- Use lexicase selection (20m)
 - Simple n' easy
 - Only useful for aggregate fitness (combination of several components)



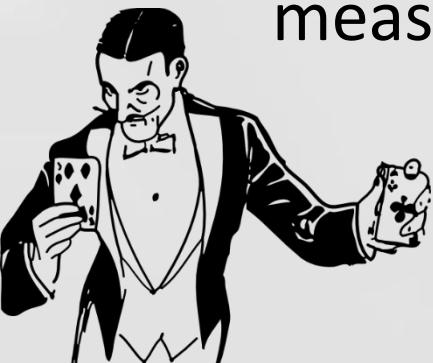
Hints and Tips

- Do you really need to promote diversity?
- Use extinction (20m)
- Use lexicase selection (20m)
- Use an island model (2h)
 - Far better than multistart (if migrations are properly handled)
 - Only useful if different experiments yield different results



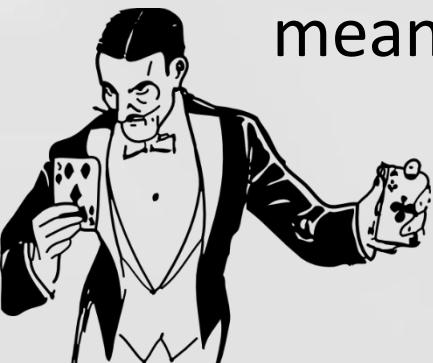
Hints and Tips

- Do you really need to promote diversity?
- Use extinction (20m)
- Use lexicase selection (20m)
- Use an island model (2h)
- Use fitness holes (20h)
 - Tweak selection operator(s)
 - Only useful if a global (and efficient) diversity measure is available



Hints and Tips

- Do you really need to promote diversity?
- Use extinction (20m)
- Use lexicase selection (20m)
- Use an island model (2h)
- Use fitness holes (20h)
- Use real niching (2-20d)
 - Only useful if the distance between genotypes is meaningful



Outline

- Generic EA
- Divergence of character in natural and artificial evolution
- Background (diversity and similarity, ...)
- Mechanisms for promoting diversity
- Hints and tips
- Conclusion

Conclusion



More Materials & Bibliography

- **MPDEA**: GECCO Workshop on Measuring and Promoting Diversity in Evolutionary Algorithms
- mpdea@polito.it
- <https://github.com/squillero/mpdea>

