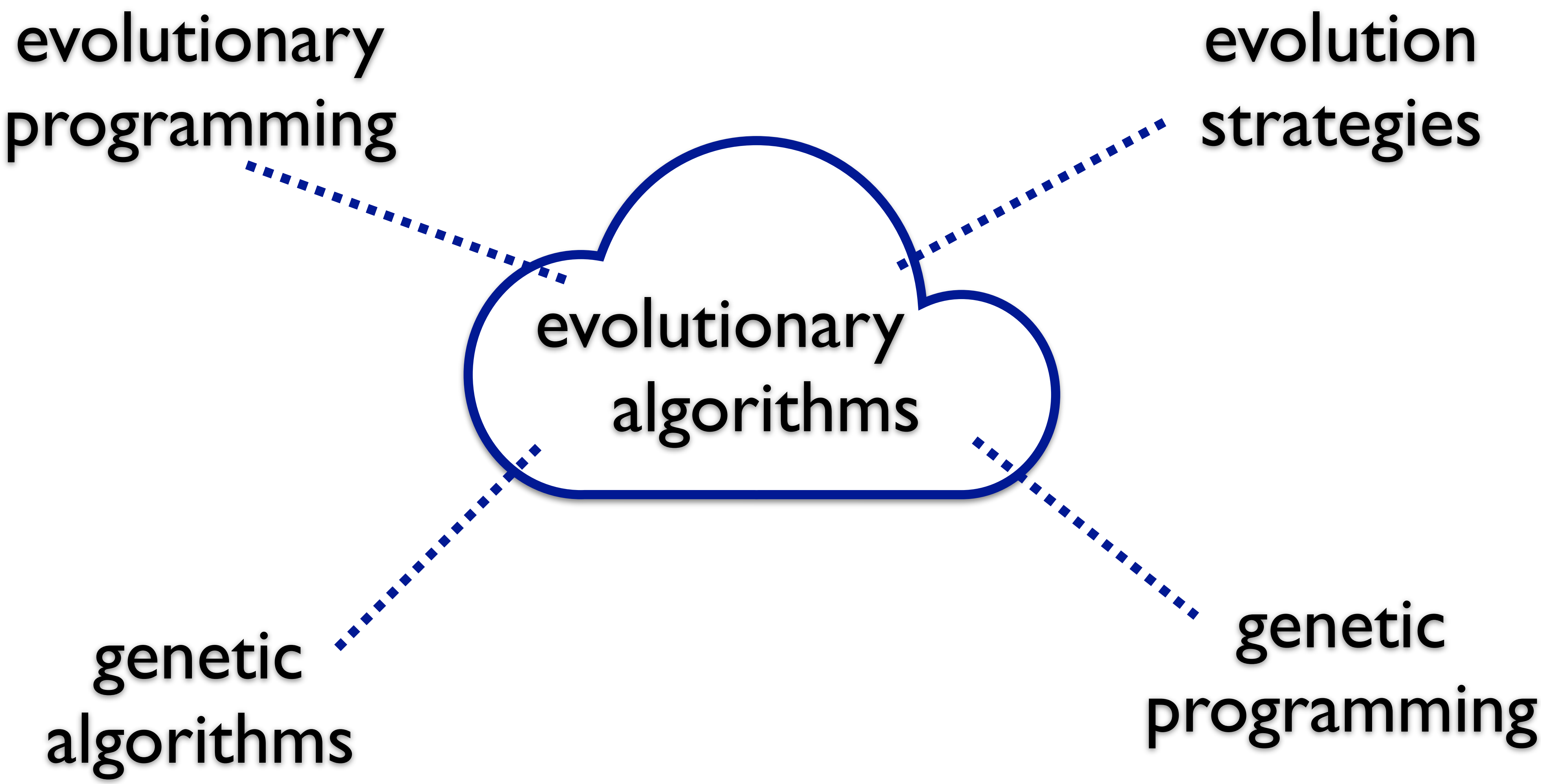
4 Evolutionary Computing Origins

Field of Evolutionary Computing



Inspiration from Biology

nature:

- consider that an environment contains a population of individuals that strive for survival and reproduction
- the fitness of each individual represents its chances of survival and multiplying
    - determined by the environment, which includes its peers

computing:

- consider using a stochastic trial-and-error style process to solve a problem, with a collection of candidate solutions
- the quality of each candidate solution is how well it solves the problem
    - and determines the chance that it will be kept and used as a seed for constructing further candidate solutions

Inspiration from Biology

Evolution		**Problem Solving**
environment	↔	problem
individual	↔	candidate solution
fitness	↔	quality

Table 2.1, Introduction to Evolutionary Computation

The Macroscopic View: Darwinian Evolution

Survival of the Fittest:

- all environments have finite resources
    - so can only support a limited number of individuals
- life forms have basic instinct/lifecycles geared towards reproduction
    - therefore some kind of selection is inevitable
- those individuals that compete for the resources most effectively have increased chance of reproduction
    - the ones that are adapted, or 'fit', the environmental conditions best

The Macroscopic View: Darwinian Evolution

Diversity Drives Change:

- phenotypic traits:
  - behaviour / physical differences that affect response to environment
  - partly determined by inheritance, partly by factors during development
  - partly as a result of random changes
  - unique to each individual
- if phenotypic traits:
  - lead to higher chances of reproduction
  - and can be inherited by offspring
- then they will tend to increase in subsequent generations, leading to new combinations of traits…

The Macroscopic View: Darwinian Evolution

Summary:

- population consists of diverse set of individuals
- combinations of traits that are better adapted tend to increase representation in population
    - individuals are "units of selection"
- variations occur through random changes yielding constant source of diversity, coupled with selection means that:
    - population is the "unit of evolution"
- and note the absence of any "guiding force"

Adaptive Landscape

Wright (1932):

- can imagine a population with n traits as existing in a $n+1$ dimensional space
    - a landscape
    - where height corresponds to fitness
- each different individual - each phenotype - represents a single point on the landscape
- population is therefore a ‘cloud’ of points, moving on the landscape over time as it evolves
    - as it *adapts*

Adaptive Landscape

- each peak represents a range of successful trait combinations
- troughs belong to less fit combinations
- so selection ‘pushes’ population up the landscape
- evolution is the process of gradual advances of the population to high-altitude areas
    - powered by variation and natural selection

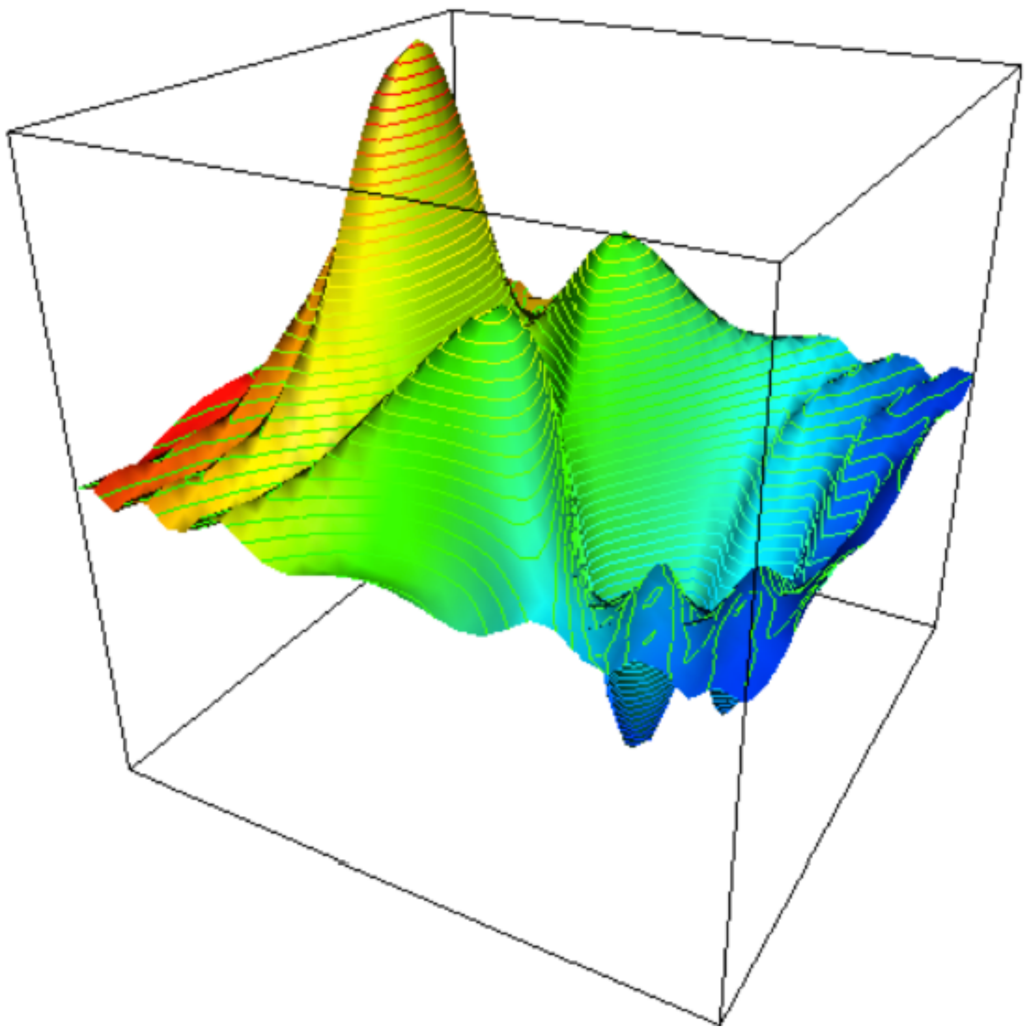


Figure 2.2, Introduction to Evolutionary Computation

Adaptive Landscape: Genetic Drift

but it's not all hill climbing:

- population has a finite size, and random choices are made in the selection and variation operators
- so it is common to observe the phenomenon of genetic drift:
    - highly fit individuals may be lost from the population
    - population may suffer from a loss of variety in some traits
- can cause the population to 'melt down' the hill, and enter low-fitness valleys
- a good thing!
    - because it it allows populations to escape from locally optima
    - giving them a chance to reach global optima

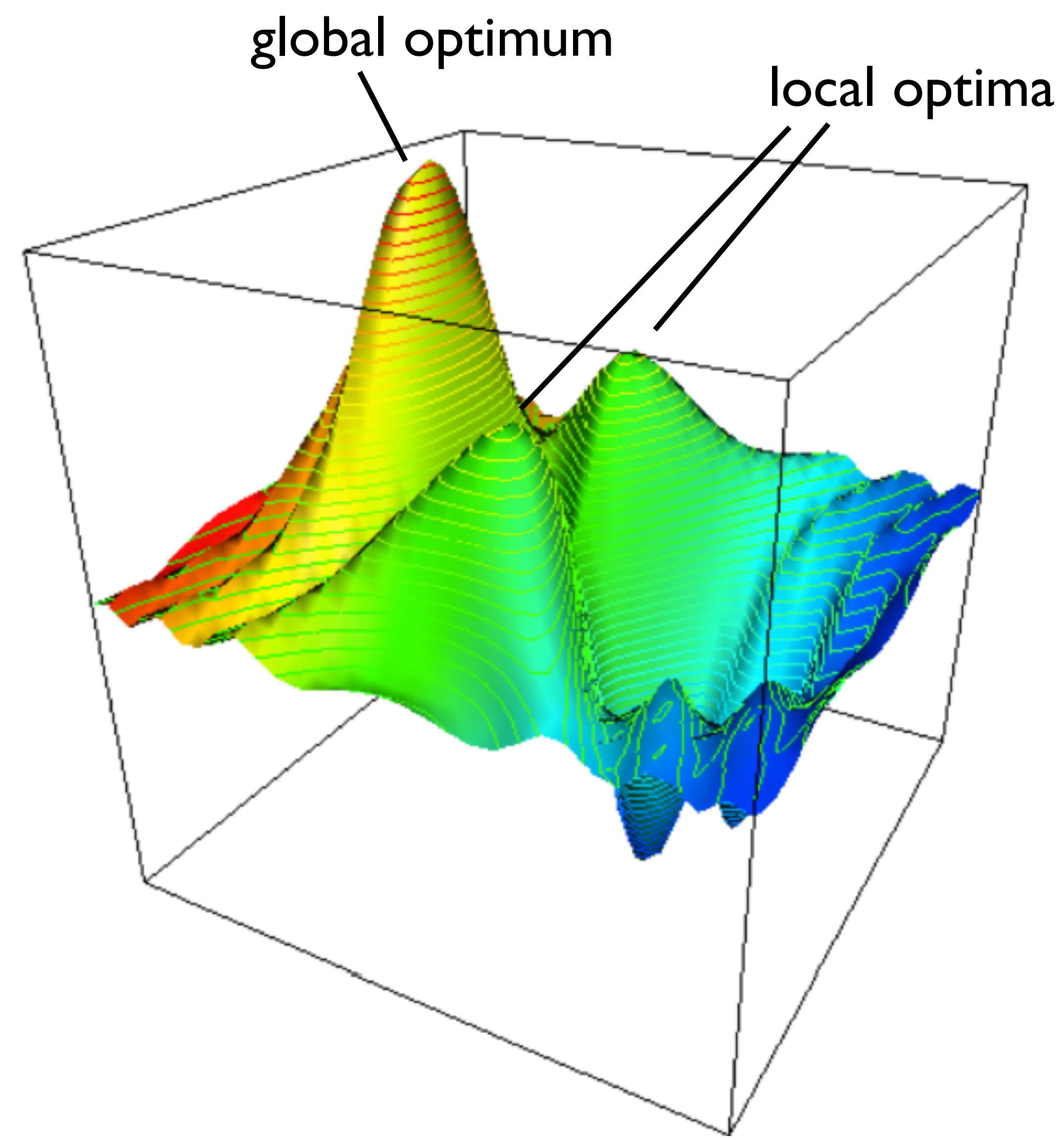


Figure 2.2, Introduction to Evolutionary Computation

The Microscopic View: Molecular Genetics

- a fundamental observation from genetics is that each individual is a dual entity:
  - its phenotypic properties (outside) are represented at a genotypic level (inside)
  - an individual's genotype encodes its phenotype
- genotype consists of genes: the functional units of inheritance
  - one gene may affect many traits (pleiotropy)
  - many genes may affect one trait (polygeny)
- small changes in the genotype lead to small changes in the organism
  - for example: height, hair colour
- allele: one of the possible values that a gene can have
  - example: if a gene in bears determines fur colour, then a polar bear has the allele that specifies the colour white

Genotype Determines Phenotype

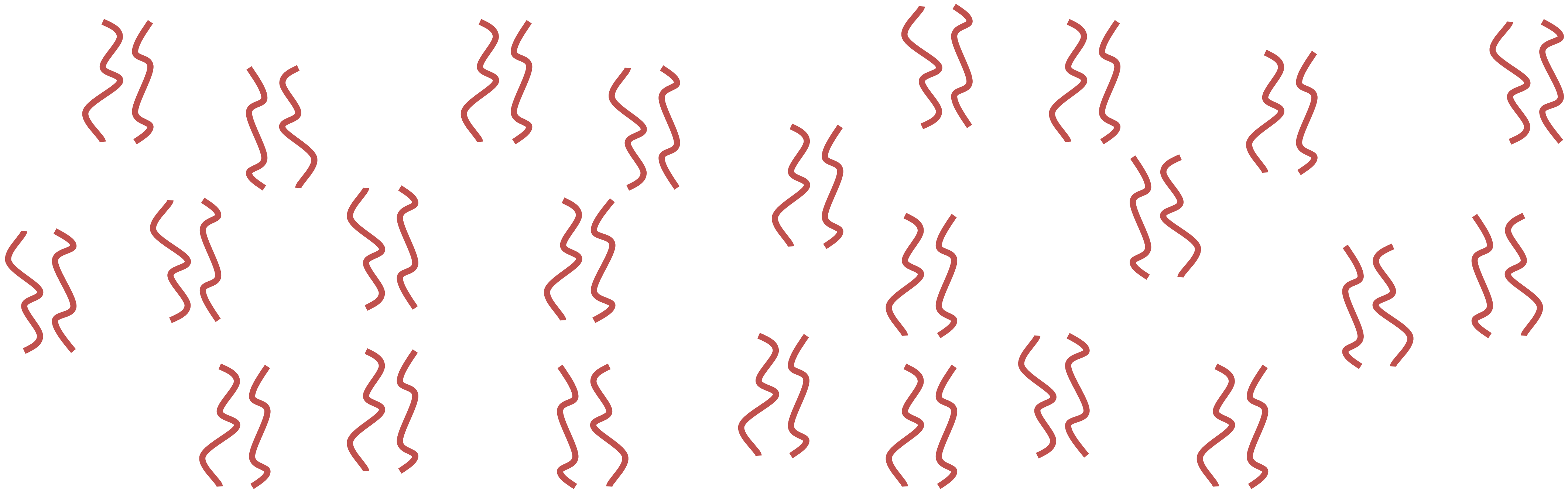
- phenotypic variations are always caused by genotypic variations
- genotypic variations are the consequences of either:
  - gene mutations
  - gene recombinations through sexual reproduction
- the genotype contains all the information necessary to build the particular phenotype

Genes and the Genome

- genes are encoded in strands of DNA called chromosomes
- most cells are diploid, holding two copies of each chromosome
  - whereas haploid cells only have one copy of each chromosome (more on this in a moment)
- the complete genetic material in an individual's genotype is called the genome
- within a species, most of the genetic material is the same

Genetics Example: Homo Sapiens

- human DNA is organised into 23 pairs of chromosomes
- together these define the physical attributes of the individual:

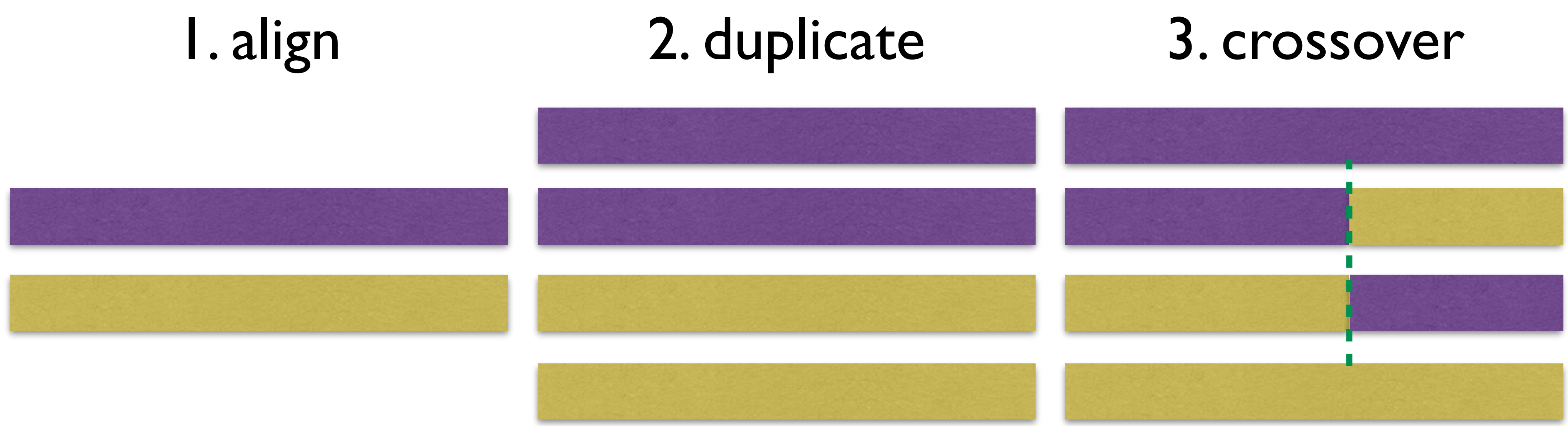


Reproductive Cells

- gametes - sperm and egg cells - contain just one copy of each of the individual chromosomes
  - so they are haploid
- they are formed by a special form of cell splitting
  - called meiosis
- during meiosis the pairs of chromosome undergo an operation called crossover

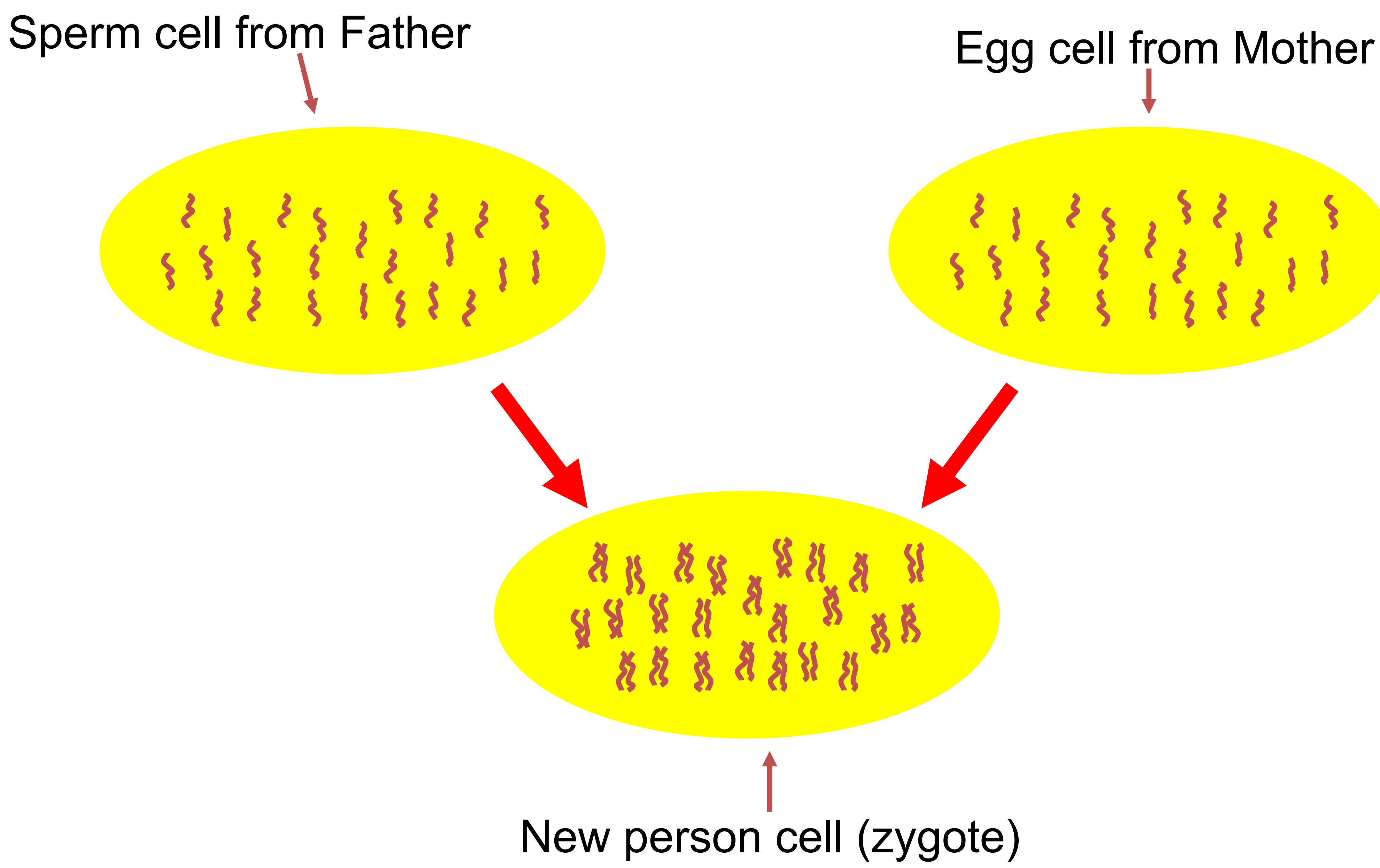
Meiosis

- chromosome pairs align and duplicate
- inner pairs link at a centromere and swap parts of themselves:



- outcome is one copy of each maternal/paternal chromosome plus two entirely new combinations
- so crossover allows new gene combinations to be created and ‘tested’
    - making it a useful technique for evolutionary computation

Fertilisation

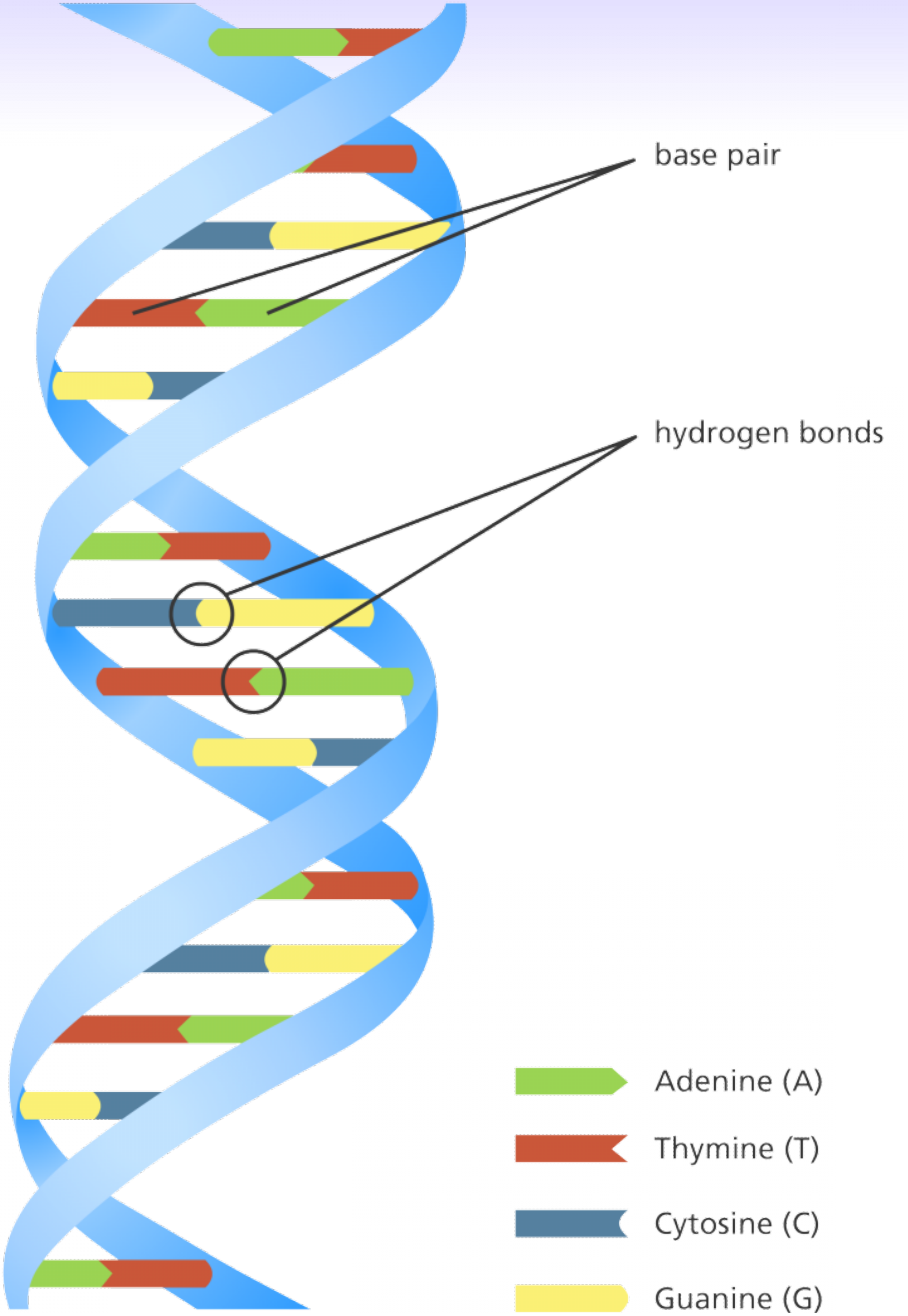


After Fertilisation

- the new zygote rapidly divides, creating many cells, all with the same genetic contents
- although all cells contain the same genes, depending on, for example where they are in the organism, they will behave differently
- this process of differential behaviour during development is called ontogenesis
- all of this uses, and is controlled by, the same mechanism for decoding the genes in DNA

Genetic Code

- all proteins in life on earth are composed of sequences built from the 20 different amino acids
- DNA is built from four nucleotides A, G, T, C, which form base pairs in a double helix spiral
    - A always bonds with T
    - G always bonds with C
- triplets of these from codons, each of which codes for a specific amino acid
- there are 64 (4^3) possible codons
- 61 of which code for the 20 amino acids
    - so there's lots of redundancy



Transcription and Translation

$$\boxed{\text{DNA}} \xrightarrow{\text{transcription}} \boxed{\text{RNA}} \xrightarrow{\text{translation}} \boxed{\text{Protein}}$$

- a central claim in molecular genetics is that the flow is one way:

  genotype $\Rightarrow$ phenotype

  phenotype $\not\Rightarrow$ genotype

Mutation

- occasionally some of the genetic material changes very slightly during this process
  - a replication error, or mutation
- so a child can have genetic material information not inherited from either parent
- the outcome of this can be:
  - catastrophic: offspring is not viable (most likely)
  - neutral: new feature does not influence fitness
  - advantageous: new feature improves fitness
- redundancy in the genetic code forms a good way of error checking

Gene Pool: another way of looking at things

- recall that the genome of the new individual is not identical to that of its parents, because of crossover and mutation…
- so genotype variations are created, which in turn translate to phenotype variations and thus are subject to selection…
- so genes are also subject to the game of survival and reproduction
- so we can view also evolution from the perspective of genes
- instead of thinking about populations of individuals, we could think about a 'gene pool'
- containing genes which compete and replicate over time, being evaluated as they reoccur in different individuals

Motivations for Evolutionary Computing

Stealing Nature's Best Ideas:

- we want automated problem solvers
- and nature is full of them!
- two most obvious candidates for inspiration:
    - the human brain
    - the evolutionary process
- designing problem solvers based on the human brain is neurocomputing
- the second option forms the basis of evolutionary computing

Motivations for Evolutionary Computing

There's Not Enough Time to Check Everything:

- the pace of change is rapid
    - leaving less and less time available for thorough problem analysis and tailored algorithm design
- problems have got more complex
- so there's a need for algorithms that are:
    - applicable to a wide range of problems
    - do not need much tailoring for specific problems, and
    - deliver good (not necessarily optimal) solutions within acceptable time
- evolutionary algorithms are well suited for this, and allow automated solution methods

Motivations for Evolutionary Computing

Curiosity:

- EC allows the opportunity to perform experiments differently from traditional biology
    - to see things as they could be
- also lets us see if we can abstract the fundamental mechanisms at play in the real world
    - although we have to be careful not to extrapolate the results too far in our conclusions
    - or to over-simplify

Reading & References

- slides largely based on, or adapted from, Chapter 2 slides for Eiben & Smith's *Introduction to Evolutionary Computing*
- while we will continue to take inspiration from nature, from now on we'll focus on how to build evolutionary algorithms using computers
- so there's no need to pick up a biology book for this course!
- but if you're interested, maybe visit the website for the Human Genome Project