Local Search & Optimization

Dr. Collin F. Lynch

AI Academy: North Carolina State University

AI Academy NC STATE UNIVERSITY

Agenda

Local Search

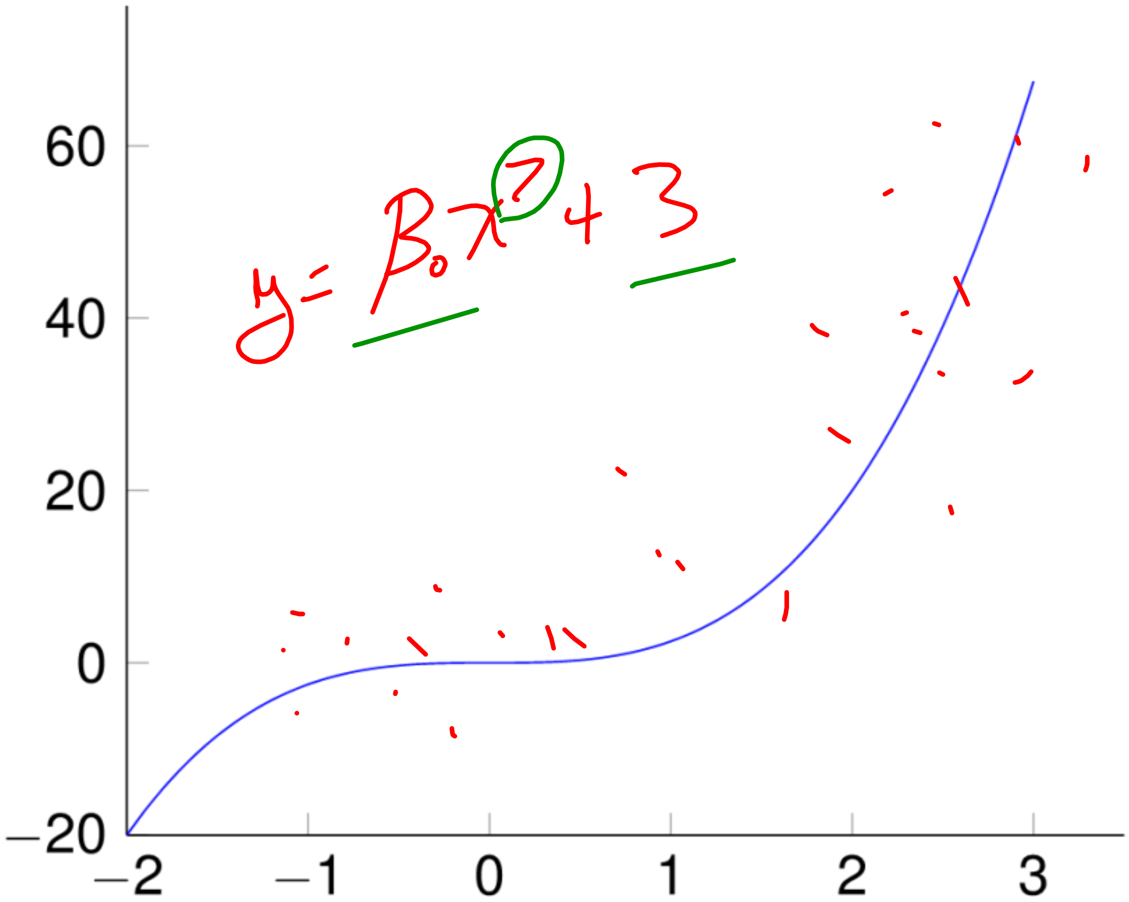
Approaches

Basics of Probability

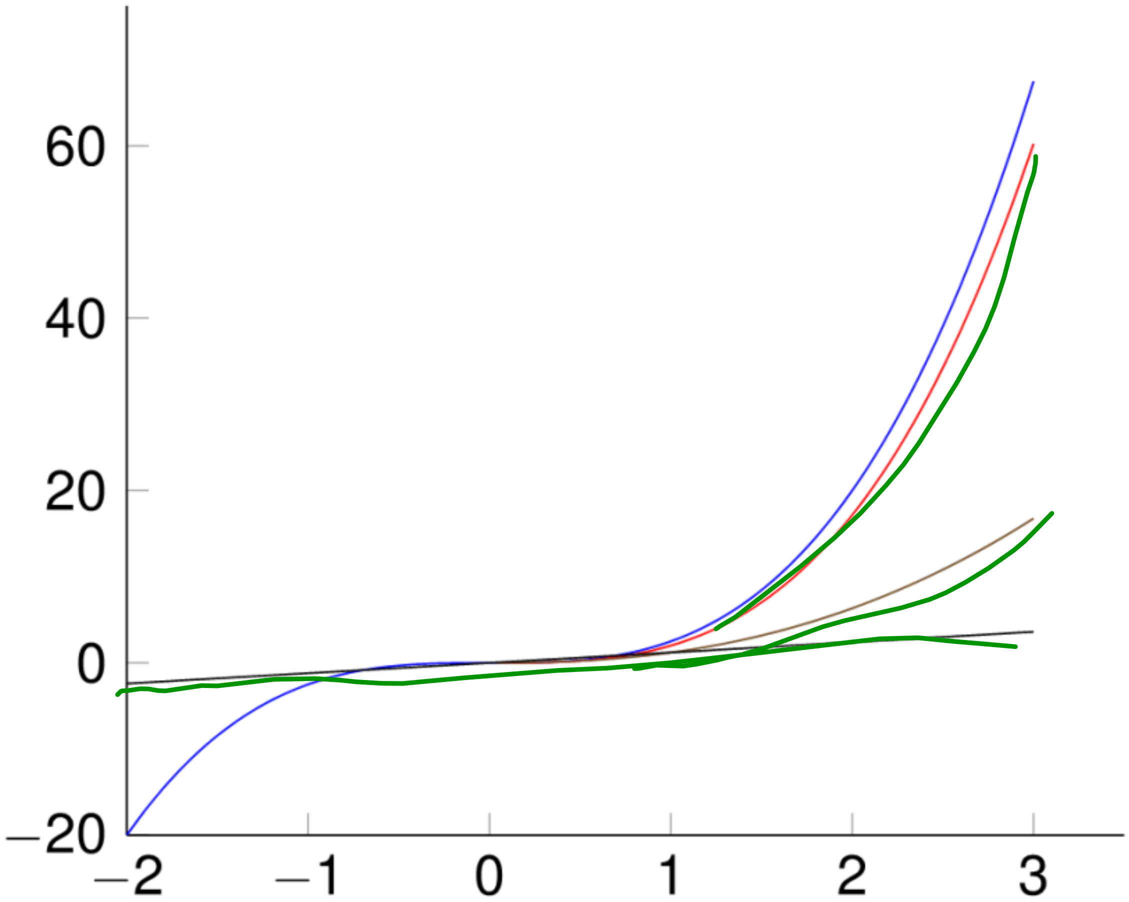
Bandits

~~Paths~~ Local Search

Function Approximation



Function Approximation (2)



Key points.

- History no longer matters.
- We use a *solution space*.
- Along with a set of possible *modifications*.
- Each solution has a potential *fitness*
- We then seek the best fit solution.

Problem Definition

Rep
Eval
Mot

1. The *solution representation*.
2. The available *actions* or *modification operations*.
3. The *Transition Model* which defines what each action does:

$$f : S_i \oplus A_j \rightarrow S_k \tag{1}$$

4. The *fitness function* (sometimes *evaluation function*).

Curve Fitting Problem

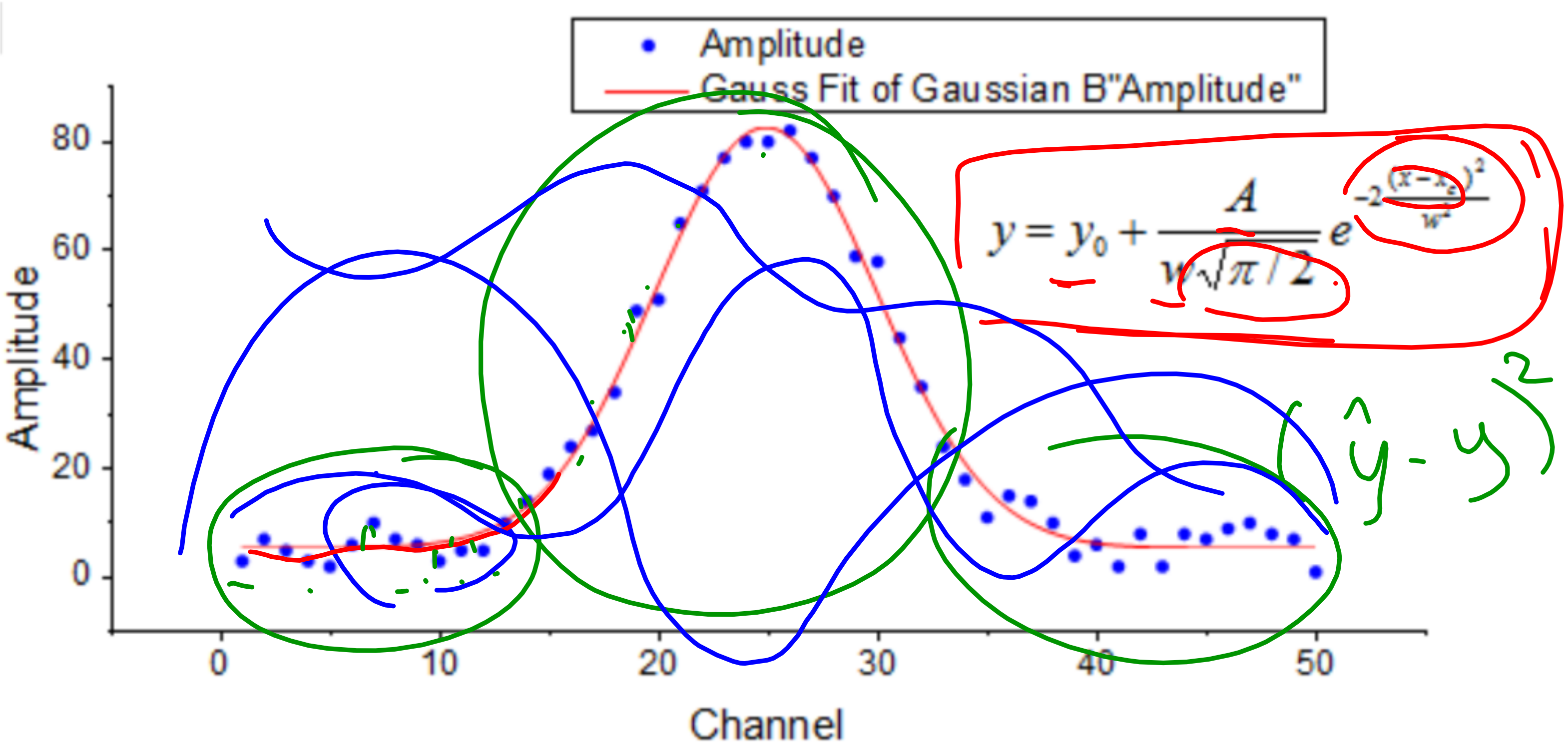


Image Credit: Originlab.com

Problem Definition

1. *Solution representation.* Equa

   $y_0 \,|\, (w) \,|\, A$

2. *modification operations.* — Increase WT

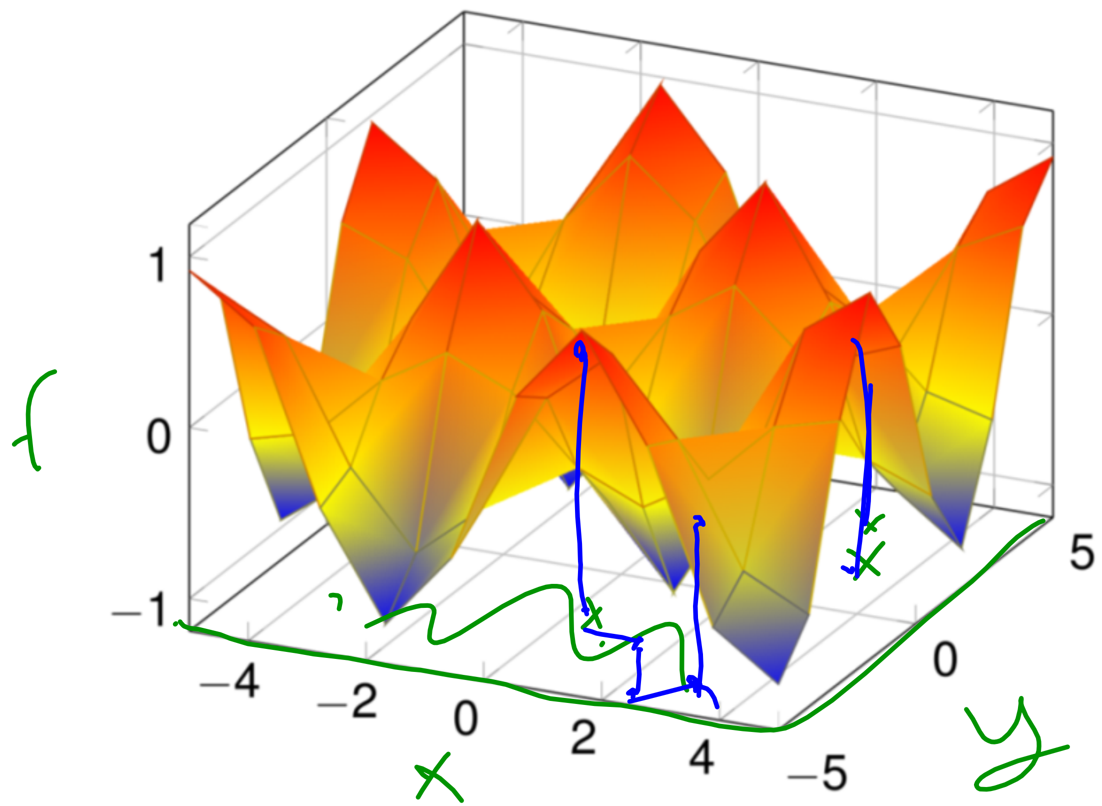
   Delete

   Add one

3. *Fitness function.*

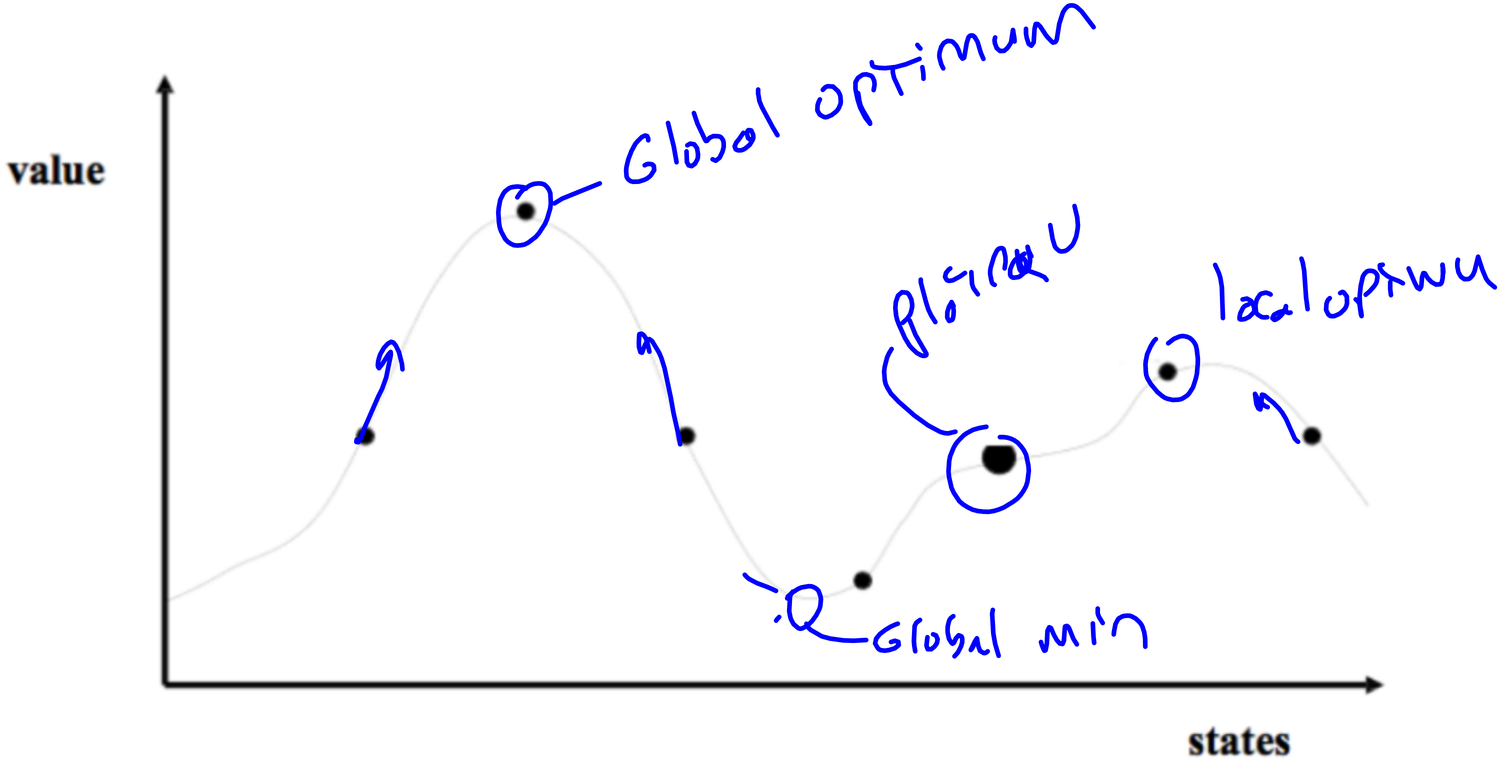
   $(\hat{y}-y)^2$

Fitness Landscape, Movement



http://tex.stackexchange.com/questions/97502/2d-colors-above-3d-surface-plot

Landscape Features



Key points

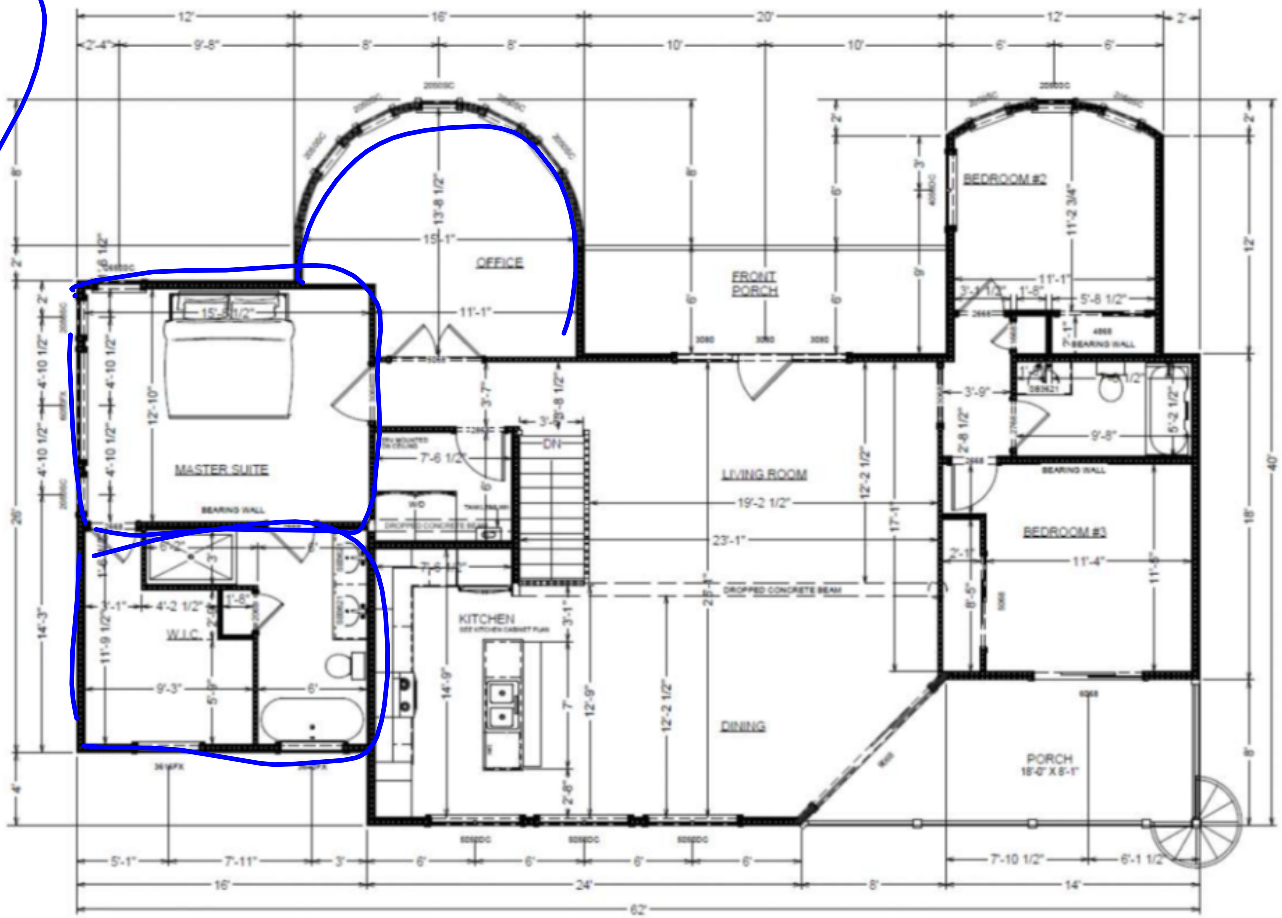
- Your definitions + algorithm, *define* the search space.
- Your evaluation function *defines* the fitness landscape.
- You want them to match.

Each other
use (client)

AI Housing

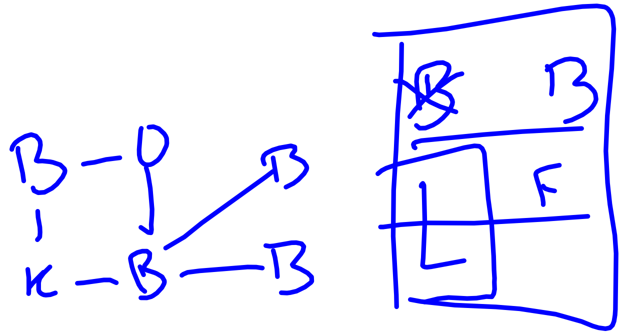
Credit: SanAntonio Express News

AI Housing



Credit: SanAntonio Express News

Problem Definition



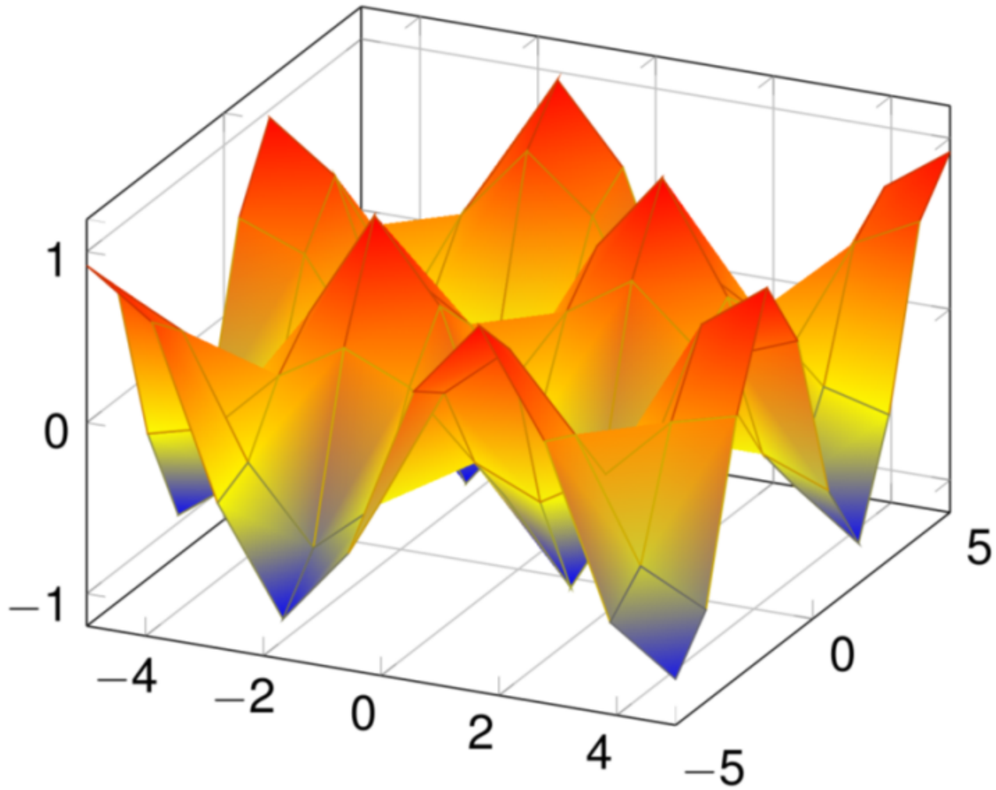
1. *Solution representation.*
2. *Modification operations.*
3. *Fitness function.*

Solu radience

Computable

Estimate Pow

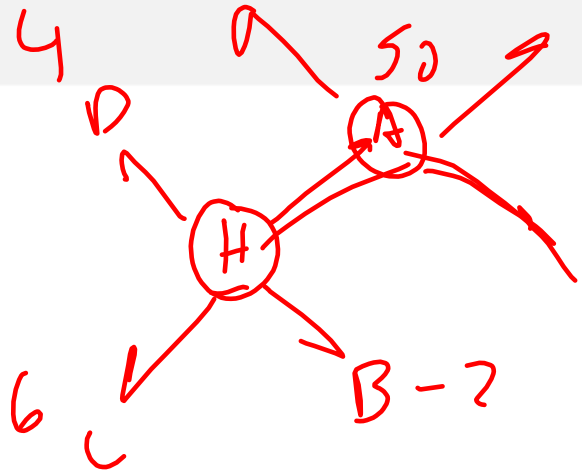
~~State~~ Search Space & Fitness Landscape

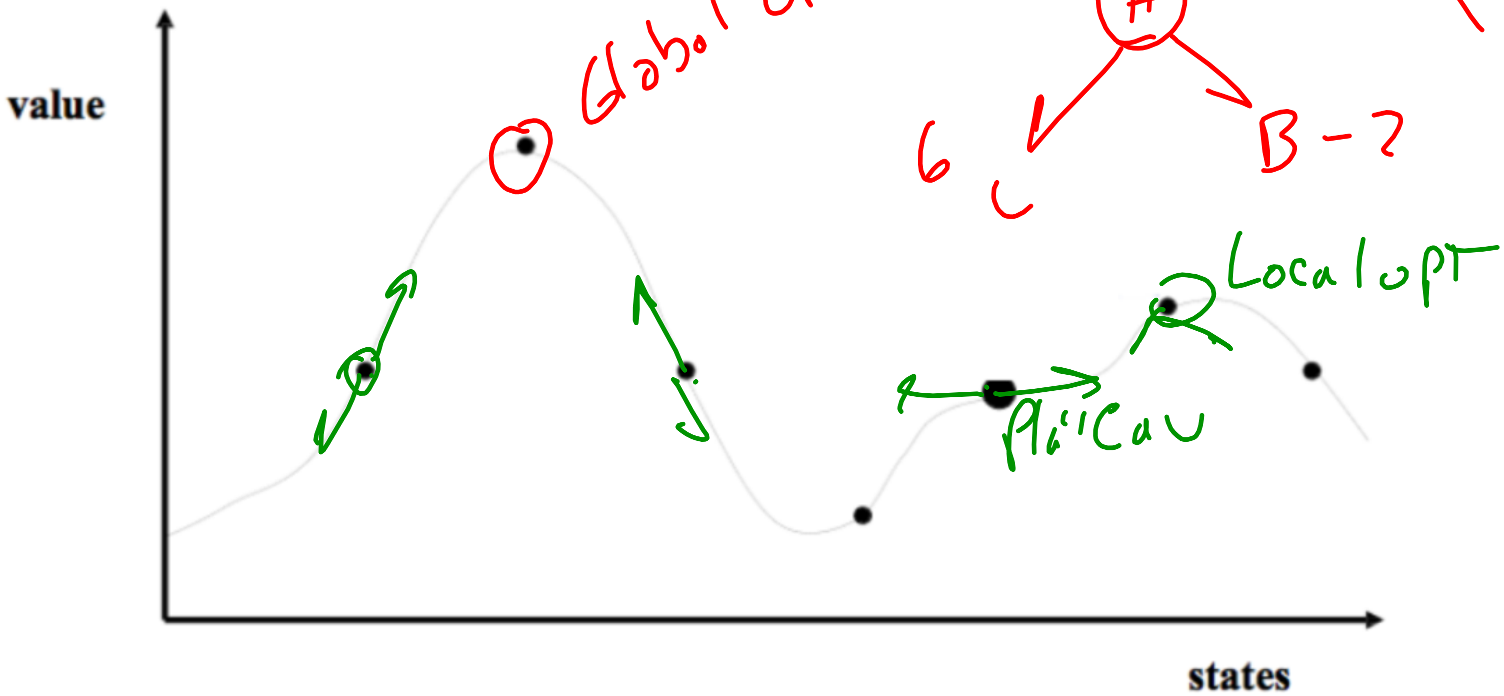


http://tex.stackexchange.com/questions/97502/2d-colors-above-3d-surface-plot

Approaches

Hill Climbing (Steepest Ascent)





Criteria

- Completeness NO
- Optimality NO
- Time Complexity — How much do you have?
- Space Complexity — Solution representation

Beam

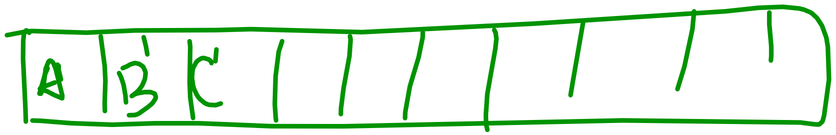
Other Hill Climbing

- Stochastic Hill-Climbing.
- Random Restart.
- Simulated Annealing: $\delta E * t$
- Beam Search.

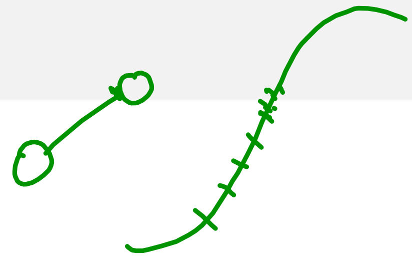
1/10 take a worse choice

Outer: Restart
Inner: Hill Clmb

Why have just one?



Continuous Gradients

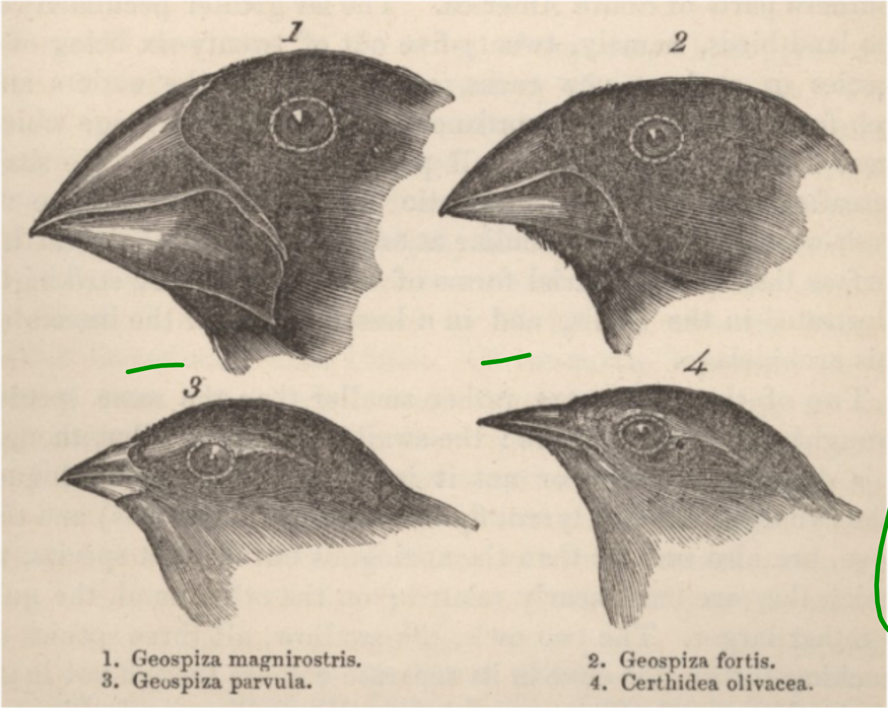


- Continuous applications allow for a true fitness landscape.
- Represent gradient as: $\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots \right)$
- Calculate (estimate) it empirically.
- Select optimal update.

This is DL
Training process

Evolutionary Computation

Submodular.



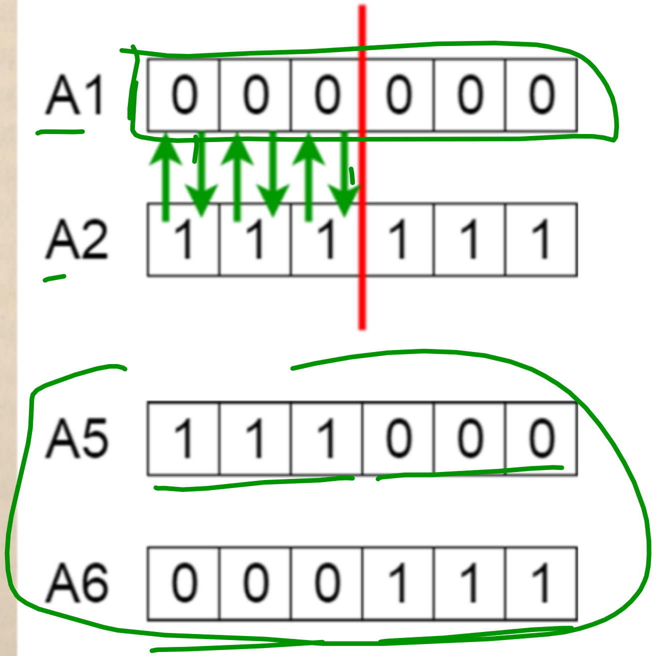


Image credit: Yale Scientific Publishers, & Vijini Mallawaarachchi

Problem Definition

AST

preservation

Exploration

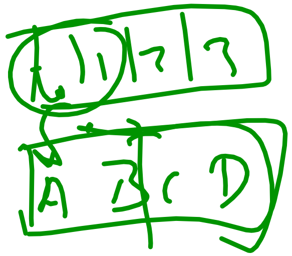
1. *Solution representation.*
2. *Genetic operations* (crossover & mutation).
3. *Selection function.*

Fitness

Tournmnt

Evolving Test Cases

- Evolve test suites for a designated block of code.
- Solution Representation: A full suite of black-box tests.
- Crossover: Exchange of tests between suites.
- Mutation: Splitting, altering IO parameters.
- Fitness Function: Code coverage and error cases.
- EvoSuite: Fraser 2018



Machine Learning is Search

Over Representations.

Basics of Probability

Foundation of Probability

Antoine Gombaud, Chevalier de Méré.

⚅⚅: one double six during 24 throws.

Credit: http://www.mon-poeme.fr

Variables Worlds and Samples

- Logical space defined by variables (e.g. $Dice_1$).
- *Sample Space (Ω):* The space of all possible worlds.
- *Probability Model:* Assignment of probabilities to states:

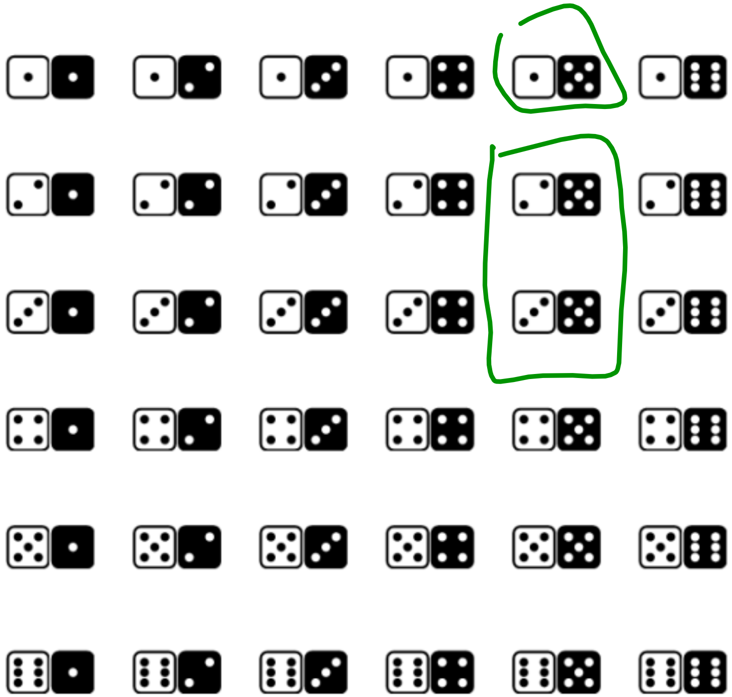
$$0 \leq p(\omega) \leq 1 \forall \omega$$
$$\sum_{\omega \in \Omega} p(\omega) = 1$$

Single Dice

D_1

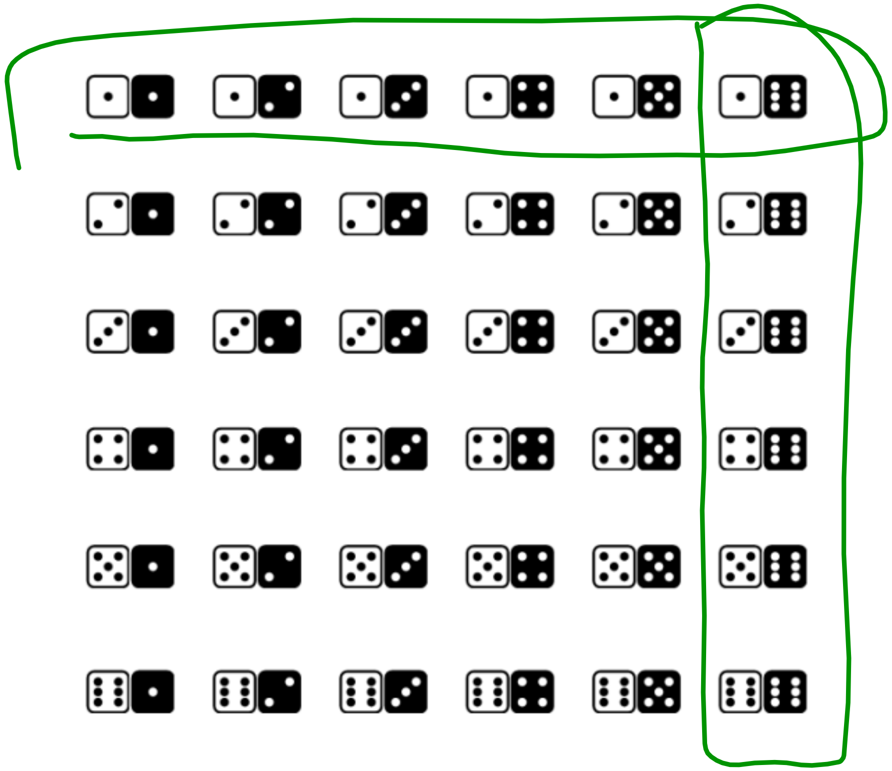
$1/6$

Joint Probability ∧



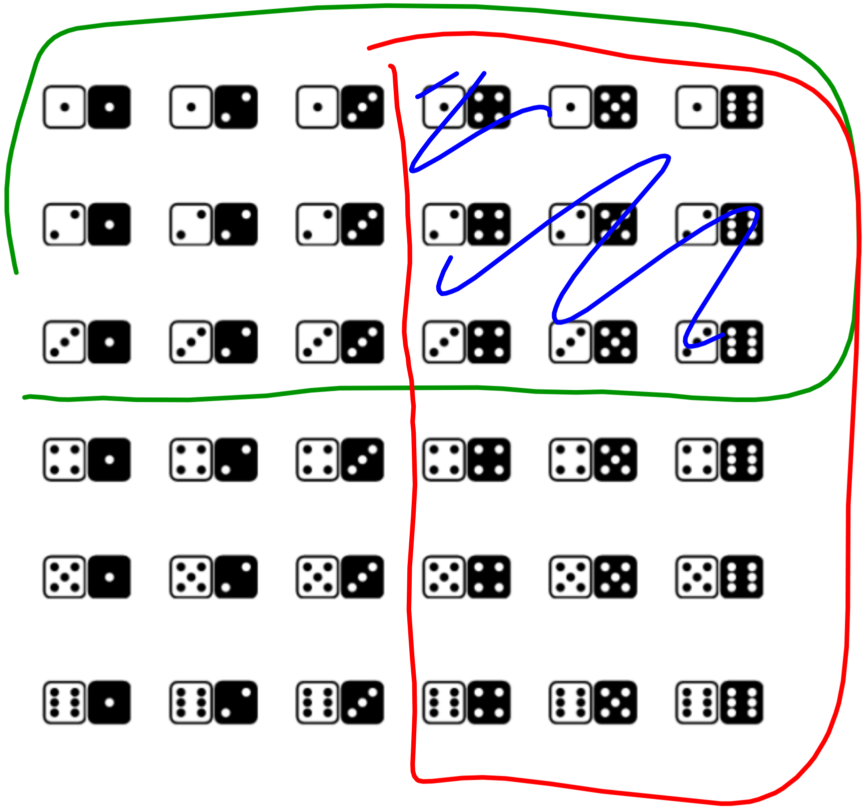
$$p(A \wedge B)$$

Joint Probability $\vee$



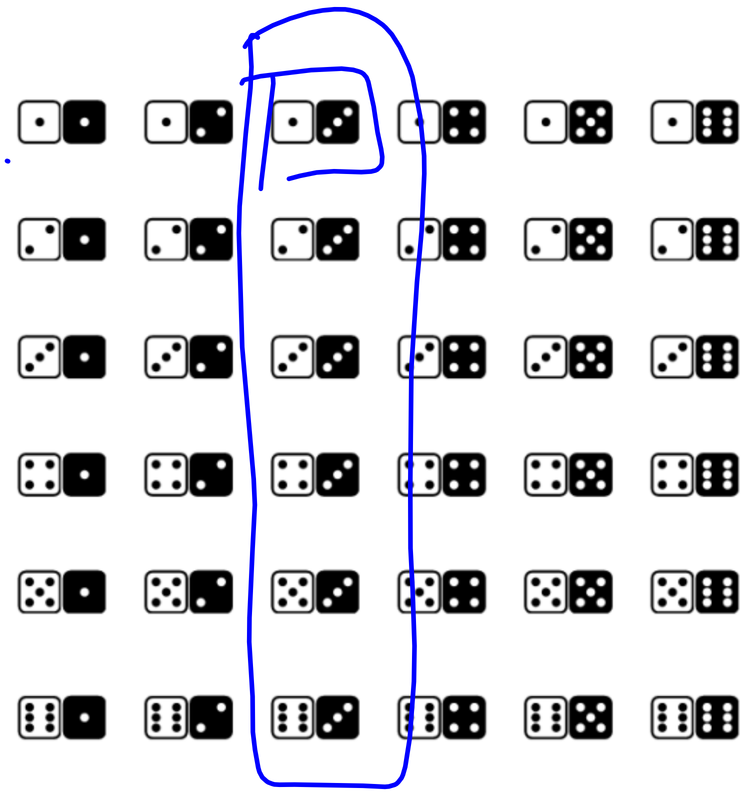
$$p(A \vee B)$$

Joint Probability Distribution



$p(D_1 < 4), p(D_2 > 2)$

Conditional Probability I



$$p(A|B) = \frac{p(A \wedge B)}{p(B)}$$

The Axioms of Probability

Name	Unconditional
Axiom	$P(A) + P(\neg A) = 1$
Fundamental Law	$P(A, B) = P(A\|B)P(B)$
Bayes Theorem	$P(B\|A)P(A) = P(A\|B)P(B)$
Independence	A, B ind. iff $P(A\|B) = P(A)$
Independence V2	A, B ind. iff $P(A, B) = P(A)P(B)$
Marginalization	$\sum_B P(A, B) = P(A)$
Expectation	$E[f] = \sum_X f(X)P(X)$

The Axioms of (Conditional) Probability

Name	Conditional
Axiom	$P(A\|C) + P(\neg A\|C) = 1$
Fundamental Law	$P(A, B\|C) = P(A\|B, C)P(B\|C)$
Bayes Theorem	$P(B\|A, C)P(A\|C) = P(A\|B, C)P(B\|C)$
Independence	A, B CI given C iff $P(A\|B, C) = P(A\|C)$
Independence V2	A, B CI given C iff $P(A, B\|C) = P(A\|C)P(B\|C)$
Marginalization	$\sum_B P(A, B\|C) = P(A\|C)$
Expectation	$EX[f\|C] = \sum_X f(X)P(X\|C)$

Probability vs. Likelihood

- We often use probability and likelihood interchangeably.
- The likelihood of B given A is $L(B|A)$
- This is the reversed conditional: $p(A|B)$

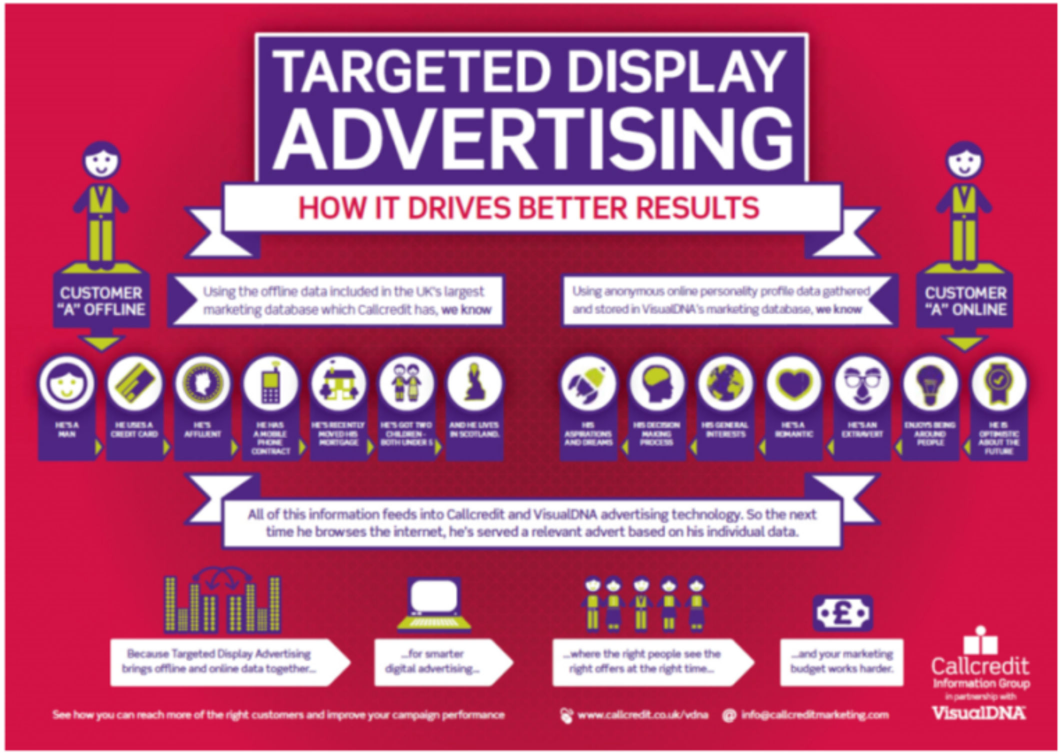
Bandits

Bandits

Exploration/Exploitation

Credit: wwwlibertygames.co.uk

Applications: Recommendation Systems



Applications: Social Networks

Core Bandits

Stationary

$$p_i = \tilde{w}_a^i (1 - \gamma) + \gamma \xi_a \tag{2}$$

- $\tilde{w}_a^i$ normalized weight for action a at time i.
- γ exploration rate.
- ξ_a uniform distribution parameter.

$$w_a^i \leftarrow \beta w_a^{i-1} + \eta r \tag{3}$$

- β decay parameter for the weight.
- η reward weight.
- r reward from current selection.

Fraser, Gordon (2018). "A Tutorial on Using and Extending the EvoSuite Search-Based Test Generator". In: *Search-Based Software Engineering*. Springer, pp. 106–130.