Solving Complex Problems with Coevolutionary Algorithms

Krzysztof Krawiec[1], Malcolm Heywood[2]
[1]Poznan University of Technology, Poland
[2]Dalhousie University, Canada

krawiec@cs.put.poznan.pl, mheywood@cs.dal.ca

http://gecco-2019.sigevo.org/

GECCO '19 Companion, July 13 – 17, Prague, Czech Republic

ACM ISBN 978-1-4503-6748-6/19/07
https://doi.org/10.1145/3319619.3323384

Instructors

- **Krzysztof Krawiec** is a Professor of Computer Science at Poznan University of Technology, Poland. His primary research areas are genetic programming and coevolutionary algorithms (CoEAs), with applications in program synthesis, modeling, image analysis, and games. Dr. Krawiec co--chaired the European Conference on Genetic Programming in 2013 and 2014, the ACM GECCO GP track in 2015 and 2016, and is an associate editor of Genetic Programming and Evolvable Machines journal. His work in the area of CoEAs includes problem decomposition using cooperative coevolution, discovery of underlying objectives in test-based problems, learning strategies for Othello using competitive CoEAs, and solving other test-based problems.

- **Malcolm Heywood** is a Professor of Computer Science at Dalhousie University, Canada. His research investigates the utility of coevolutionary methods under non-stationary environments (e.g., streaming data and financial applications), and uses coevolution to facilitate the discovery of agents for reinforcement learning tasks in games such as the Atari Learning Environment, VizDoom and Dota 2. Dr. Heywood is a member of the editorial board for Genetic Programming and Evolvable Machines (Springer). He was a track co-chair for the GECCO GP track in 2014 and a co-chair for European Conference on Genetic Programming in 2015 and 2016.

Agenda

- I. Introduction
- II. Competitive coevolution
  - Core concepts
  - One-population competitive coevolution
  - Two-population competitive coevolution
  - Advanced topics
- III. Cooperative coevolution
  - Core concepts
  - Case study: Evolving arbitrary sized teams
  - Case study: Non-stationary streams
  - Case study: Diversity maintenance and policy reuse
  - Case study: Multi-task learning under Arcade Learning Environment
  - Case study: Combining Competitive and Cooperative coevolution
- IV. Closing remarks

I. Introduction

Canonical assumptions made by EA

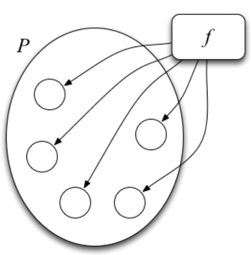
- An **absolute measure** of fitness is **available** and **computable**.
  - ‘Complete’ definition of task / environment
- Solutions are (more or less) **monolithic**.
  - Each individual encodes a **complete solution** to a problem
  - Tasks are not explicitly ***decomposed***.
- Coevolutionary algorithms (CoEA) **revise these assumptions**.

What is a coevolutionary algorithm?

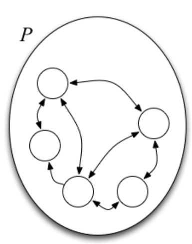
- A variant of EC where **fitness function mandates the individuals to engage into direct interactions**.
  - Fitness cannot be computed for isolated individuals.
- Formally:
  - Evolutionary algorithm (**EA**): $f: X \rightarrow E$
  - Coevolutionary algorithm (**CoEA**): $f: X_1 \times X_2 \times \ldots \times X_n \rightarrow E$, where E is an evaluation codomain (typically R)
  - **Interaction** = a tuple from $X_1 \times X_2 \times \ldots \times X_n$

EA vs. CoEA

EA
Absolute measure of fitness f available and computable for each individual separately.



CoEA
Search gradient can be obtained only by letting individuals interact. Exact fitness may be not computable.



Consequences

- Individuals' performances depend on each other (fitness is contextual)
- The solution of a problem can be:
  - An **element** of X_i (as in an EA)
    - Typical for ***competitive CoEA*** (with exceptions)
    - Key questions: What to evolve against? Who is the best teacher?
  - A **combination of elements** from X_is
    - Typical for ***cooperative CoEA*** (with exceptions)
    - Key questions: How to encourage cooperation? Divide and conquer.
- Pertains to so-called solution concepts, see later
- Remember: individual ≠ solution

What is it good for?

- CoEAs lend themselves conveniently to a few classes of problems of theoretical and practical interest.
- **Competitive CoEAs**: test-based problems, games, interactive domains
  - Example: individual=game strategy, fitness=expected game outcome
- **Cooperative CoEAs**: problem decomposition, modularity, credit assignment
  - Example: individual=a rule in a classifier, fitness=overall accuracy of the classifier
- Class of problems: co-search, co-optimization, generalized optimization (Wolpert and Macready 2005)

Other characteristics of CoEAs

- Operate under **incomplete information** (uncertainty)
- Focus on evaluation and **interaction schemes** (less so on search operators)
- Individuals often maintained in **several populations**.
- Biological analogs:
  - No global, static fitness function in Nature
  - Nature does not optimize for anything; EAs do.
  - Individual's fitness results from its interactions with environment, including other individuals of the same species

Measuring progress: Subjective vs. objective fitness

- **Subjective fitness**: *f* calculated using the currently available elements of X_is (a sample)
  - Typically those available in the current population,
  - Example: average game outcome against the opponents from the current population
- **Objective fitness**: *f* calculated with the elements chosen in a principled manner.
  Examples:
  - Average game outcome against all possible opponents
  - Game outcome against a human-crafted opponent.

II.1. Competitive coevolution

Class of problems tackled by competitive CoEAs

- Interactive domains
  - Sets of **individuals** (**entities***)
  - **Interaction function** (payoff function)
    $g: X_1 \times X_2 \times \ldots \times X_n \to R$
  - When n=2, the second argument is an opponent.
- Note: g alone does not define the search goal.
- What is the solution to the problem?

(*) Sometimes, but not always, identified with candidate solutions

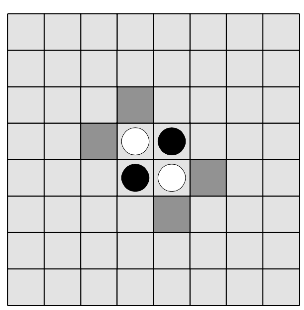
- Solution concept (cf. Ficici 2004, Popovici et al. 2012):
  - Criterion specifying whether a potential solution
    - is better than another solution (in **co-optimization**),
    - is solution to a problem (in **co-search**)
- Most popular SC: Maximization of Expected Utility (MEU):
  $f_o(x_1) = \mathrm{E}[\ g(x_1,x_2)\ ]$
  - A.k.a. generalization performance (Chong et al. 2008)
- Competitive CoEAs realize knowledge-free approach to solving problems posed in interactive domains.

Subjective fitness

- Challenge: calculation of f_o is often computationally infeasible.
  - Example: game of Othello: game tree complexity 10^{58}
  - Number of unique strategies typically much higher
- Solutions:
  - 1. **Fix the set of opponents.**
    - For instance, well-performing known opponents (e.g., handcrafted by humans)
    - Strong bias, limited generalization
  - 2. **Draw the opponents at random**
    - What is the 'right' distribution of opponents?
    - Drawing uniformly in the genotypic space does not result in desired (e.g., uniform) distribution of skills/capabilities
  - 3. **Competitive coevolution**

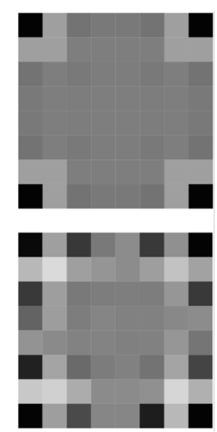
Example: Game of Othello

- Two-player, perfect-information, turn-based, zero-sum game
  - Still unsolved
  - Sudden changes of game state possible
- Strategy = candidate solution
- Competitive CoEA approach:
  - Evolve board evaluation function $b()$
  - Use it in one-ply search: simulate all legal single moves from the current state and choose the one that maximizes b.
- Popular representations of board evaluation functions: weighted piece counter and n-tuples



Weighted Piece Counter (WPC)

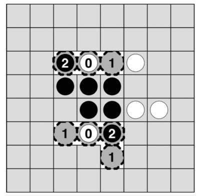
- Single linear neuron with 64 weights: $b(s) = \Sigma_i\ w_i s_i$
- Top: handcrafted Othello WPC board evaluation function (*standard WPC heuristics*)
- Bottom: a function evolved using one-population competitive CoEA, hybridized with temporal difference learning (TDL) (Szubert, Jaśkowski, Krawiec 2009)

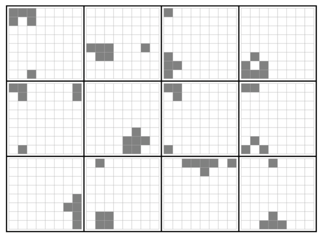


N-tuple networks

(Browning 1959, Lucas 1997)

- Combinatorial network with lookup tables holding all combinations for (usually randomly selected) subsets of (usually adjacent) board locations
- 3^n weights for a single n-tuple for tri-state boards (for Othello: empty, black, white)
- Top: 3-tuple and 4-tuple for base-3 numbers (white, empty, black):
  - $2*3^2 + 0*3^1 + 1*3^0 = 19$
  - $1*3^3 + 0*3^2 + 2*3^1 + 1*3^0 = 34$
- Bottom: Examples of CTDL co-evolved n-tuples (Szubert, Jaśkowski, Krawiec, 2013)





One-population competitive CoEA

- The simplest setup to approach MEU problems.
  - Applicable when $X_1 = X_2 = ... = X_n = X$
  - E.g. symmetric games
  - Usually: $f_s(x) = \Sigma_{x' \in X'} g(x,x')$, where X' is a sample drawn from current population P
- Interaction = single game (symmetric games) or two games (asymmetric games)
- Interaction schemes:
  - **Round-robin**: $n(n-1)/2$ interactions ($X' = P \setminus \{x\}$)
  - ***k*-random opponents**: kn interactions ($|X'| = k$)
  - **Single-elimination tournament** (SET): n interactions
    - Pair the individuals at random. Winners pass to the next stage. Fitness is the stage reached in the tournament.

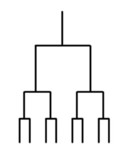
Highlights of one-pop competitive CoEAs

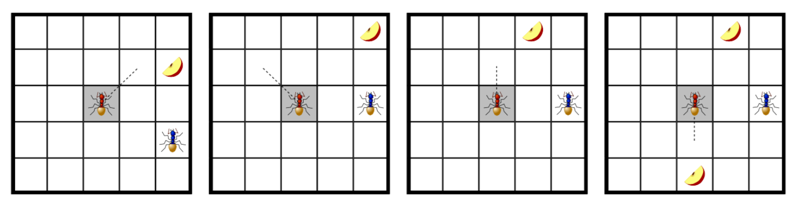
- Iterated Prisoner's Dilemma (Axelrod 1987)
- Backgammon (Pollack & Blair 1998)
- Checkers (Samuel 1959, Fogel 2002)
- NERO, Blackjack, Pong, Small-board Go, Tetris, …

Fitnessless Coevolution for Ant Wars

(Jaśkowski, Krawiec, Wieloch 2008)

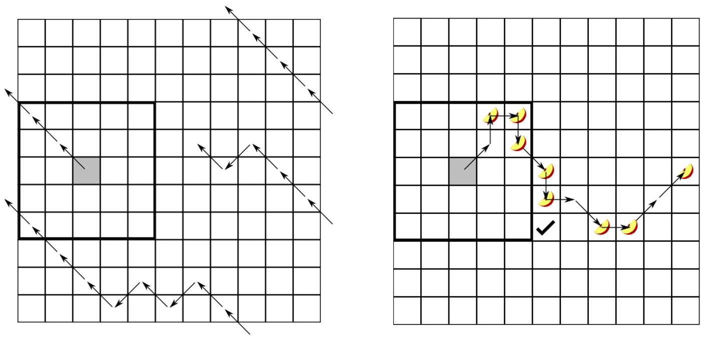
- FC: Pick k individuals at random. Run a SET on them and return the winner.
- Evolved the winner of the Ant Wars GECCO'08 contest
  - Two-player partially observable game
  - Agents (ants) see only a 5x5 fragment of the toroidal 11x11 board
  - The goal: collect more food pellets than the opponent.
  - Strategy representation: stateful GP program (intra-game memory)





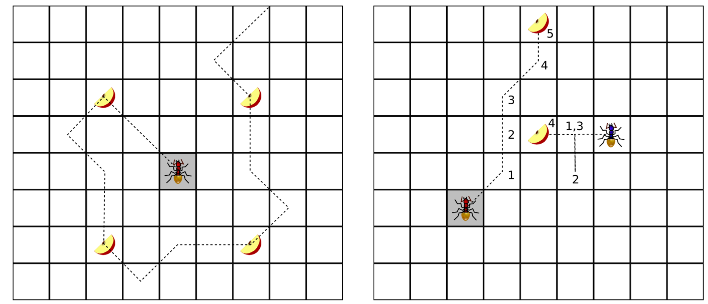
Example: Ant Wars

Complex behaviors emerged: **systematic search, rational choice of trajectories**, …



Example: Ant Wars

… **memorizing locations of food pellets, opponent avoidance, pseudo-suicide,** …



- Online demo: http://www.cs.put.poznan.pl/kkrawiec/antwars/

Digression: Importance of transitivity

- Fitnessless Coevolution is not equivalent to fitness-driven one-population coevolution if there are cycles in interactions in between individuals (Jaśkowski, Krawiec, Wieloch 2008)
- Example: Tic-tac-toe strategies A, B, C: place a mark in the numbered locations if free, otherwise in the location marked by asterisk (*)

A

1	2	3
		*

B

3		
2		
1		*

C

		1
	2	
3		*

- A beats B, and B beats C. But A does not beat C, just the opposite.
- Tic-tac-toe is **intransitive**.
- **No scalar fitness function** can model this (can realize only complete orders).

One-pop competitive CoEAs as self-learning

- Individuals create **search gradient** for each other.
  - A form of (population-level) **self-learning**
  - Can be seen as an analog to **self-play** in RL (individual-level)
- Q: Is this sufficient to guarantee progress?
- A: No.
  Coevolutionary pathologies are lurking out there.

Coevolutionary pathologies

- **Cycling**: evolution keeps rediscovering the same solutions
  - Particularly likely if game is intransitive.
- **Disengagement**: opponents are either trivial or way too difficult to beat.
- **Overspecialization** (**focusing**): mastering the skills of beating some opponents while neglecting the others.
- **Forgetting**: opponents defeated in the past turn out to be difficult again.
- See review and rigorous analysis in (Ficici 2004)
- Main causes:
  - No access to objective fitness
  - Population responsible for **both search and providing search gradient** for itself

Coevolutionary archive competitive CoEAs (one-population)

Archive = a container storing well-performing individuals, maintained alongside the population.

- Provides **long-term memory**
- Builds **search gradient**
- Prevents some **pathologies**
- Maintains **diversity** and **progress**

Archives help maintaining **historic progress** (Miconi 2009)

- Not necessarily progress in the global, objective sense.

How it works:

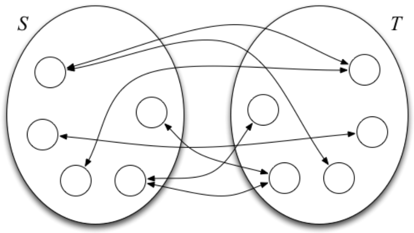
- Search algorithm submits some individuals to the archive
- Archive accepts some of them
- Individuals in population interact with peers **and** archival individuals
- Outcomes of interactions augment the fitness
- Simplest archive: best-so-far individual
- Hall of fame (Rosin & Belew, 1997)
  - Stores all best-of-generation individuals found so far
  - Population members play against each other and against the opponents from HoF

II.2. Two-population competitive CoEAs

Two-population competitive CoEAs

- One-pop competitive CoEA: Population responsible for both searching for good solutions and providing search gradient for itself.
  - Why not separate these functions?
- Two-pop competitive CoEAs: maintain separate populations of:
  - **candidate solutions** $S \subset X_1$ – intended to **solve the problem**
  - **tests** $T \subset X_2$ – **provide <u>only</u> search gradient** for the individuals in S
- Applicable in symmetric ($X_1 = X_2$) and asymmetric setting ($X_1 \neq X_2$)

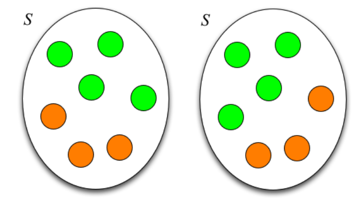
Two-population competitive CoEA



- Typical **interaction scheme**: all-to-all
- *S* and *T* co-evolve in parallel
- No transfer of individuals between *S* and *T*

How to evaluate the tests?

- Individuals in *S* rewarded for ***performing*** (aim at maximizing EU).
- **How to reward the tests** in *T*? Maximize EU as well?
  - Pathologies likely
  - Tests should be neither too easy nor to hard for the individuals in *S*
- Idea: reward tests for ***informing***, e.g.:
  - **Distinctions**: for every pair of distinguished solutions
  - **Informativeness**: for unique partitioning of *S*
  - Hybrids (e.g., with EU)



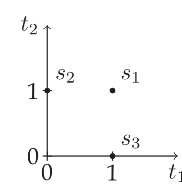
Test-based problems

- With two populations, the tests can be conceptually different from candidate solutions.
- **Test-based problem** (*S*, *T*, *G*, *Q*) (Popovici et al., 2012)
  - *G* – payoff matrix
  - *Q* – solution quality function
- Examples:
  - **Asymmetric games** (strategies vs. opponents)
    - E.g., tic-tac-toe, Othello,
  - **Control problems** (controllers vs. initial conditions)
    - Pole balancing, car control, etc.
  - **Learning from examples** (hypotheses vs. examples)
  - **Program synthesis** with GP (programs vs. tests)
  - In general: **co-optimization** and **co-search**
- Also applicable in symmetric settings

Pareto-coevolution

(Ficici and Pollack, 2001; Noble and Watson, 2001)

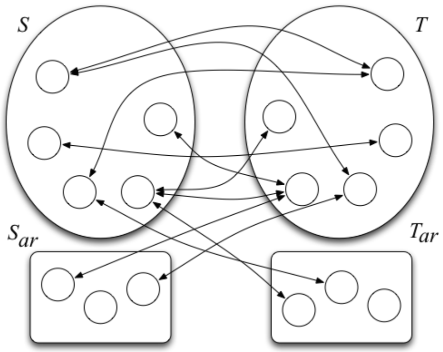
- Each test considered as a **separate objective.**
- Transforms a test-based problem into a **multiobjective optimization problem** (or many-objective one).
- Example:
  - s_1 solves both tests t_1 and t_2
  - s_2 solves only t_2
  - s_3 solves only t_1
  - s_1 **dominates** both s_2 and s_3



- Problem: **large number of tests** (and thus objectives).
- **Sparse** dominance relation.

Coevolutionary archives (two-pop)

- General scheme: individuals are **submitted** to archive and get **accepted or rejected** by it.
- **Separate archives** for solutions and tests



Coevolutionary archive algorithms (two-pop)

- **Iterated Pareto-Coevolutionary Archive**, IPCA (de Jong 2004)
  - A new solution s is added to S_{ar} if no $s' \in S_{ar}$ dominates s. In that case:
    - All $s'' \in S_{ar}$ dominated by s are removed from S_{ar}
    - The test t that made it possible for s to be added to S_{ar} is added to T_{ar}
  - Guarantees monotonous progress
  - Unlimited-size archive
  - Tests provide for **distinctions** between individuals
- **Layered Pareto-Coevolutionary Algorithm**, LAPCA (de Jong 2004)
  - Merges the current archive and the submitted elements and builds a Pareto ranking of solutions
  - The first k layers of the ranking remain in S_{ar}, the remaining ones are discarded
  - T_{ar} keeps the tests that support Pareto dominance in S_{ar}
  - No guarantee of monotonous progress, but (somehow) controllable size
- IPCA and LAPCA perform well only on small, usually artificial problems.

Coevolutionary archives

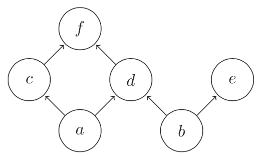
- Maintaining archives can be **costly**
  - Many interactions required to check if a solution should be added
- Mitigation: MaxSolve (De Jong 2005), for MEU solution concept
  - Keep in S_{ar} up to n solutions that solve the most tests (at least one), and in T_{ar} all tests that a solved by at least one $s \in S_{ar}$
  - [Behaviorally] duplicate tests are discarded
  - Monotonic: will not miss solutions that increase the number of solved tests
- When overhead of maintaining an archive counted in, non-archived algorithms can be equally efficient.
- Other types of archives (Jaśkowski & Krawiec 2010)
- Related concepts: **ideal evaluation** and **complete evaluation set** (E. de Jong and Pollack 2004)
  - The set of tests that preserves dominance relation between the solutions in S
  - Determining the minimal complete evaluation set is NP hard (Jaśkowski & Krawiec 2011)

II.3. Advanced topics in competitive coevolution

(selection)

Coordinate systems

- An interaction matrix defines a **dominance relation**
- Dominance relation defines a **partial order** in the set of individuals ⇒ partially ordered set, ***poset***

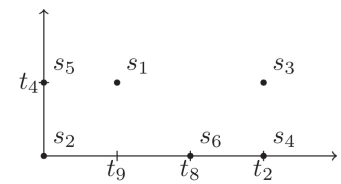


- A poset can be 'stretched' along multiple **dimensions** (**underlying dimensions**).
- Dimensions form a **coordinate system** (Bucci et al. 2004):
  - Axis = ordered list of tests (the most popular formulation)

Coordinate system: an example

- The game: Nim-1-3
  - Players in turns take sticks from two piles of size 1 and 3.
- Total of 144 strategies,
  - but only 6 behaviorally unique for the first player (*S*), and 9 for the second player (*T*).
- **Minimal** coordinate system
  - Some tests not needed to reproduce the dominance relation
- Game dimension: 2

	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9
s_1	1		1	1		1	1		1
s_2									
s_3	1	1	1	1	1	1	1	1	1
s_4	1	1	1				1	1	1
s_5	1			1			1		
s_6							1	1	1

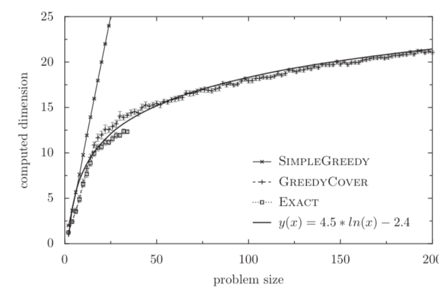


Coordinate systems: some results

- Can accelerate convergence and/or guarantee progress: Dimension Extraction Coevolutionary Algorithm, DECA (de Jong and Bucci 2006)
- Reveal the **internal structure of a problem** and relate to **problem difficulty**
- Hypothesis: dimensionality of coordinate system is a yardstick of **problem difficulty**
- The set of all tests forms the **complete evaluation set** (de Jong & Pollack 2004)
- Game dimension = **width of the poset** (Jaśkowski & Krawiec 2011)
- The number of underlying objectives for an abstract problem seems to be limited by a logarithm of the number of tests.

Problems with exact coordinate systems

- Problem dimension may be **underestimated** when only samples of *S* and *T* are used.
- Finding minimal CS for a problem is **NP-hard** (Jaśkowski & Krawiec 2011)
- Heuristics exist but **overestimate** the number of dimensions
- Nontrivial test-based problems have **very high dimensionality**
- Q: Can we efficiently **'approximate' the underlying dimensions**?



Heuristic discovery of underlying objectives

- Idea:
  - Construct **efficiently** approximate underlying objectives from the information available at the given stage of search process
  - Use the derived objectives in multiobjective EA setting
- **Derived objectives** rather than **underlying objectives**
  - **Approximate** (do not reproduce the original dominance)
  - **Transient** (depend on the current populations)
- Technical means: **clustering** of tests

Heuristic discovery of underlying objectives

(Krawiec & Liskowski 2015, Liskowski & Krawiec 2016)

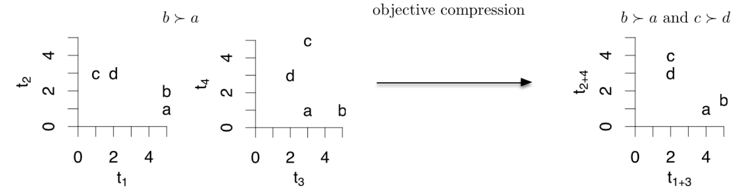
$S = \{a, b, c, d\}$ $T = \{t_1, t_2, t_3, t_4\}$

G	t_1	t_2	t_3	t_4
a	5	1	3	1
b	5	2	5	1
c	1	3	3	5
d	2	3	2	3

objective compression

derived objectives t_{1+3} and t_{2+4}

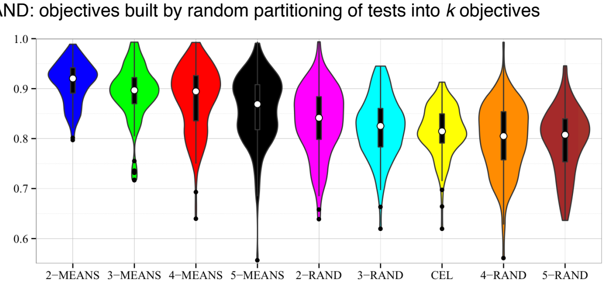
G'	t_{1+3}	t_{2+4}
a	4	1
b	5	1.5
c	2	4
d	2	3



Upside: denser dominance relation. Downside: 'false positive' dominance possible

Heuristic discovery of underlying objectives

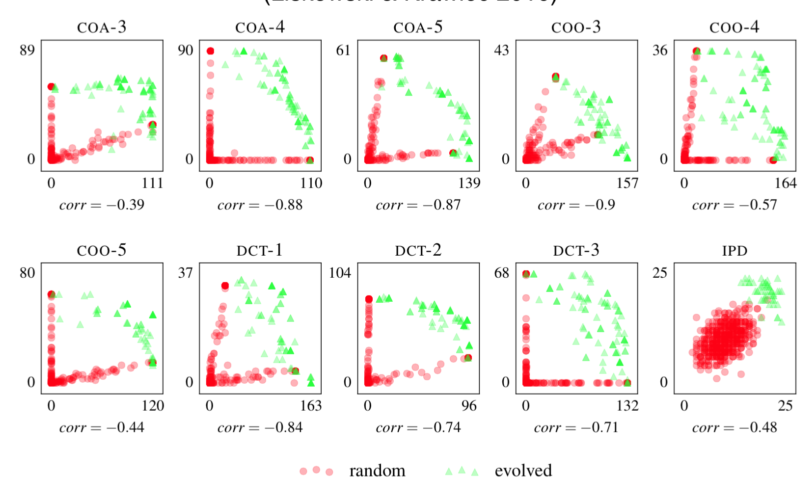
- Results for 9-choice iterated prisoner's dilemma, IPD (MEU)
  - *k*-MEANS: *k* objectives derived using *k*-means clustering algorithm
  - *k*-RAND: objectives built by random partitioning of tests into *k* objectives



- Applied also in non-coevolutionary setting with GP, with *k* adjusted automatically (Krawiec & Liskowski 2015). Better than GP and RAND, comparable to IFS.

Heuristic discovery of underlying objectives

(Liskowski & Krawiec 2016)



Genetic Programming: Program synthesis as a test-based problem

- In GP, programs are evaluated by confronting them with (a sample of) tests
  - S = population of candidate programs
  - T = population of tests (fitness cases)
- Simple variant: Pairwise Comparison of Hypotheses (Krawiec 2001)
- Fully-fledged coevolutionary approach: (Arcuri & Yao 2014)
  - Synthesis from formal specification (precondition + postcondition)
  - Co-evolving sets of unit tests in T alongside with programs in S
  - Strongly-typed GP
  - Tested on nontrivial benchmarks: MaxValue, AllEqual, TriangleClassification, Swap, Order, Sorting and Media
  - Better than random sampling of tests (particularly when using specialized sub-populations corresponding to parts of formal specification)
- Related: collecting test cases from program verification in spec-based GP (Krawiec, Bladek, Swan 2017)

Genetic Programming: Alternative definitions of underlying objectives

- Non-negative matrix factorization (NMF) allows decomposing the interaction matrix G into a pair of matrices W, H, where the columns of W can be interpreted as underlying (derived) objectives: DOF (Liskowski, Krawiec 2016)

$$G = \begin{array}{c c} & \begin{array}{cccc} t_1 & t_2 & t_3 & t_4 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{pmatrix} 2 & 2 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix} \end{array} \xrightarrow[rank(G)=2]{NMF \quad k=2} W \times H = \begin{array}{c c} & \begin{array}{cc} f_1 & f_2 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \\ s_3 \end{array} & \begin{pmatrix} 0.70 & 2.05 \\ 0.73 & 0.66 \\ 0.35 & 1.02 \end{pmatrix} \end{array} \times \begin{array}{c c} & \begin{array}{cccc} t_1 & t_2 & t_3 & t_4 \end{array} \\ \begin{array}{c} f_1 \\ f_2 \end{array} & \begin{pmatrix} 0.70 & 0.70 & 2.70 & 2.70 \\ 0.74 & 0.74 & 0.06 & 0.06 \end{pmatrix} \end{array}$$

$p_1 \succ p_2, \quad p_2 \succ p_3, \quad p_1 \succ p_3$

$p_1 \succ p_3$

	s_1	s_2	s_3
fitness:	4	2	0
ranking:	1	2	3

	s_1	s_2	s_3
ranking f_1	2	1	3
ranking f_2	1	3	2
average ranking	1.5	2	2.5

- Empirical evidence for DOF outperforming standard GP
- NMF can be applied also to *sparse* matrices: SFIMX (Liskowski, Krawiec 2016):
  1) Perform only a fraction of interactions in G.
  2) Use NMF to restore the complete G and so define a *surrogate fitness*.
- Related: Neural Estimation of Interaction Outcomes (Liskowski, Krawiec, Wieloch 2018)

Hybridization with RL

- CoEAs are **generate-and-test** techniques (like EA)
  - In contrast, **gradient-based** methods provide 'directed' corrections/updates of parameters
  - Can be more efficient in high-dimensional problems
  - Complementary: CoEAs learn slower than TDL but eventually outperform it (Lucas & Runarsson 2006)
- **Coevolutionary Temporal Difference Learning**, CTDL (Krawiec & Szubert 2011, Szubert et al. 2013)
  - Interleave one-population coevolution (with round-robin) with TD(0)
  - CoEA picks the 'right' opponents, TDL tunes the solutions in a self-play mode
  - CoEA modifies the topology of n-tuples. TDL only affects the weights.
- A form of **memetic algorithm** (genetic local search) (Moscato 1989): individuals' interactions with the environment influence their genotypes (Lamarckian evolution).
- Related to: **adversary reinforcement learning**

Hybridization with RL

- Othello, *n*-tuples (Szubert, Jaśkowski, Krawiec 2013)
  - Compared also to ETDL= EA+TD(0)
- Othello Evaluation Function League
  - Ranked according to average performance against *standard heuristic* WPC (handcrafted strategy; moves partially randomized) (as of 2011)
  - http://algoval.essex.ac.uk:8080/othello/html/Othello.html
- ETDL better on predefined opponent (heuristic WPC)
- CTDL produces more versatile players

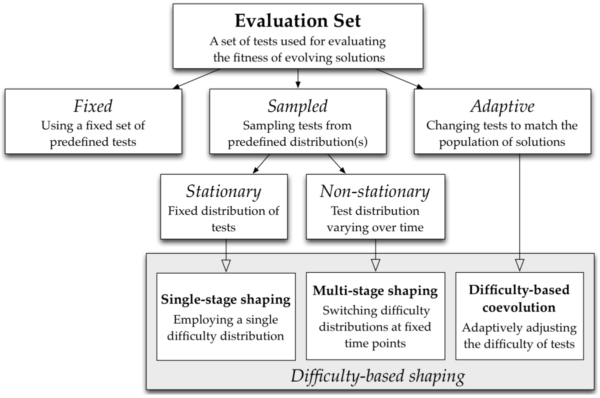
Othello League Ranking

Name	Size	Played	Won	Drawn	Lost
epTDLmpx_12x6	12×6	100	89	1	10
prb_nt30_001	30×6	100	84	0	16
prb_nt15_001	15×6	100	83	3	14
epTDLxover	12×6	100	81	4	15
t15x6x8	15×6	100	79	3	18
SelfPlay15	12×6	100	77	0	23
tz278_2	278×2	100	76	3	21
Nash70	12×6	100	72	4	24
x30x6x8	30×6	100	71	4	25
pruned-pairs-56t	56×2	100	71	1	28

Coevolutionary shaping

- Shaping = key concept in behavioral psychology (Skinner 1938)
  - **Expose the learner to a series of training episodes of gradually increasing difficulty.**
  - Motivation: Tasks can be too difficult to learn autonomously.
  - Example: To train a pigeon to strike a ball, first reward looking at it, then approaching it, and only then striking the ball with the beak.
- Used with success in Reinforcement Learning, e.g. pole balancing (Selfridge 1986)
  - Simplified version of tasks generated by relaxing/parameterizing the original one
  - E.g. change the length of the pole, increase the mass, etc.
- Related to: **incremental evolution, staged evolution, environmental complexification**
- Requires human intervention (decide how to relax the tasks, order them, etc.)

Coevolutionary shaping



Bottom line: Coevolutionary shaping works as well as manual shaping, but requires **less parameter tuning** (Szubert 2014, Szubert et al. 2013)

Competitive Coevolution: Key take-home messages

- A competitive CoEA **can** guide itself towards the optimum, even though it does not have access to objective fitness.
  - However, this can be ineffective due to **pathologies**.
- Archives (and populations of tests in two-pop coevolution) form long-term memory and **accumulate knowledge** about a problem.
- Coordinate systems and **underlying objectives** are examples of alternative **search drivers**.
  - Aim at widening the 'evaluation bottleneck' and making search algorithm **better-informed.**
- CoEAs are particularly effective for **adversarial problems**.
  - Many problems of practical interests can be posed in this way.

Not covered in this tutorial

- Measuring and visualizing progress (e.g., CIAO plots)
- Artificial problems: number games. Strategies represented as vectors of n elements.
  - **Compare-on-all:** A solution wins if it is better on all elements
  - **Compare-on-one:** a test picks a dimension at random; the solution wins if it's greater on that dimension
- Other solution concepts (Ficici 2004, Poppovici et al. 2011)
  - Simultaneous maximization of all outcomes, Nash equilibrium, Pareto-optimal set, Algorithms: (Ficici 2004) and review in (de Jong 2005)
- Deciding upon the final outcome of a CoEA: "output mechanism" (Popovici and Winston 2015)
- Random Sampling Evolutionary Algorithm (Chong et al. 2008) - no true coevolution, but hard to beat using competitive CoEAs.
- Coevolutionary free lunches (Wolpert & Macready 2005; Service and Tauritz 2008; Popovici and Winston 2015)
- Hybridization with CMA-ES (Jaśkowski & Szubert, 2015)
- In-depth analysis of relations between test-based co-optimization and supervised learning (Popovici, 2017)

III. Cooperative Coevolution

Cooperative Coevolution

- **Answers the question:**
  - How to encourage collaboration?
- **Metaphor:**
  - Divide and conquer!
- **Why (is it useful?): Promoting modularity / reuse**
  - additional clarity in: (relative to a monolithic solution)
    - credit assignment
      - search space projected into multiple smaller search spaces
      - agents do not need to solve all the task
    - solution transparency
    - capacity to react to changes (Simon's parable of the two watch makers)
- **Fitness: who to credit for what?**
  - generalist pathology:
    - individuals rewarded for maximizing the number of collaborations
    - stable / mediocre solutions rather than optimal solutions

A Metaphor…

- "species [individuals] represent *solution components*. Each individual forms a part of a complete solution but *need not represent anything meaningful on its own*. The components are evolved by measuring their contribution to complete solutions and recombining those that are most beneficial to solving the task." [Gomez et al., (2008)]
- **Central questions**
  - How to:
    - How to **compose** a candidate solution (team)
    - **Distinguish** between credit to the team versus that to team members
    - Learn **context**
    - Maintain **diversity**
    - Continuously **adapt/complexify**

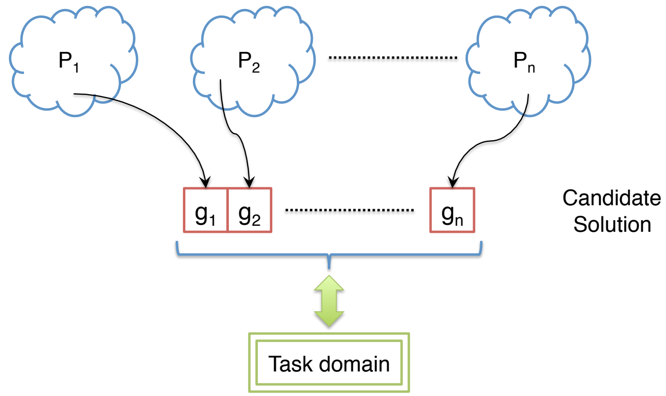
Cooperative Coevolution for complex systems : Some milestones

- **Neural Networks**
  - Moriarty, Miikkulainen (1998)
  - Potter & de Jong (2000)
  - Gomez et al. (2008)
- **Genetic Programming**
  - Krzystof & Bhanu (2006, 2007)
  - Thomason & Soule (2007), Rubini et al. (2009)
  - Lichodzijewski & Heywood (2008)
  - Wu & Banzhaf (2011)
- **Formulating fitness functions**
  - Panait et al. (2006, 2008)
  - Agogino & Tumar (2008), Knudson & Tumar (2010)
- **Non-stationary tasks**
  - Agogino & Tumar (2008)
  - Vahdat et al, (2015)
- **Heterogeneity versus homogeneity**
  - Agogino & Tumar (2008)
  - Waibel et al. (2009)
  - Gomes et al. (2018)
- **Diversity maintenance**
  - Lichodzijewski et al. (2011)
  - Doucette et al. (2012)
  - Kelly & Heywood (2014, 2018a)
  - Gomes et al. (2014, 2017)
- **Reinforcement Learning (RL)**
  - Luke et al. (1997)
  - Moriarty & Miikkulainen (1998)
  - Gomez et al. (2008)
  - Agogino & Tumar (2008)
  - Knudson & Tumar (2010)
  - Rubini et al. (2009)
  - Doucette et al. (2012)
- **Visual Reinforcement Learning**
  - Kelly & Heywood (2017a,17b, 18b)
  - Smith & Heywood (2018, 19)
- **External memory for Coop. Coev.**
  - Smith & Heywood (2019a, 19b)

Cooperative Coevolution: An architecture

(Potter & De Jong, 2000)

Prior decomposition of the solution into ‘n’ independent populations (species)



Biased and Lenient cooperation

(Panait et al., 2006), (Panait et al., 2008)

Biased cooperation

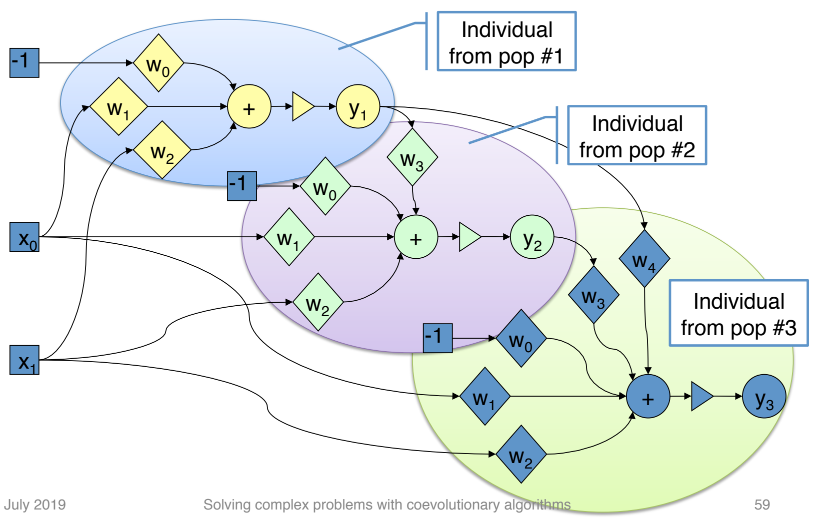
- Consider team versus individual fitness
  - Individuals receive avg. of fitness from teams
  - Promotes generalists
  - Hitchhiking
- Recommend defining individual fitness as
  - an *optimal* team of collaborators
  - Not clear how an *optimal* collaborator set is found in the general case

Lenient cooperation

- Individual fitness
  - $\mathrm{MAX}_{i \text{ in team}}$ (team$_i$ fitness)
  - Hitchhiking still exists
  - Is hitchhiking all negative?
    - Enables individuals to find their niche
    - Provides a memory of previous / alternative policies

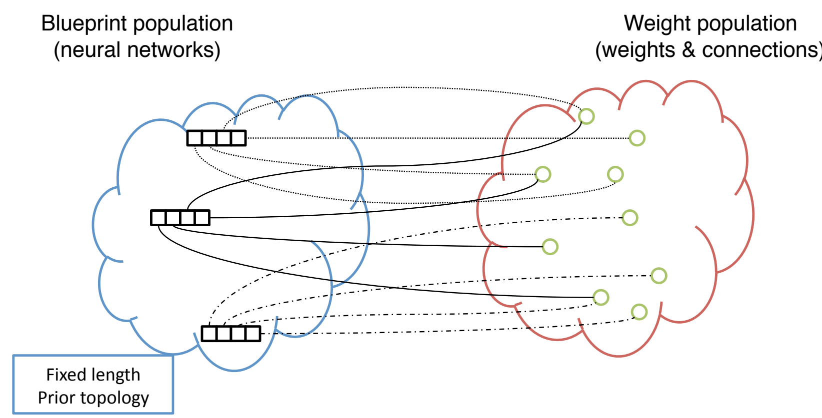
Coevolving a cascade network

(Potter & De Jong, 2000)



SANE with blueprints

(Moriarty & Miikkulainen, 1998)



Difference evaluation functions

(Agogino & Tumar, 2008), (Knudson & Tumar, 2010), (Codly & Tumar, 2012)

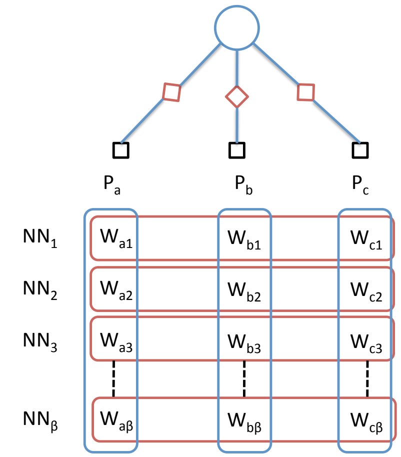
- **Global fitness**
  - Performance of entire collective
  - Difficult to identify the contribution from each agent
- **Local fitness**
  - Performance of single agent
  - Difficult to encourage non-overlapping collective behaviours
- **Difference evaluation function (D_i)**
  - Explicitly estimate value added by agent 'i'
  - Global fitness needs to be locally 'decomposable'
  - Agents assigned w.r.t. physical locality to distributed sub-tasks
  - Form of 'spatial embedding'
- **D_i formulation**
  - $D_i = G(s) - G(s_{-i} + C_i)$
- G(s)
  - G(•) is the global evaluation function
  - 's' state of the system
- s_{-i}
  - States for which agent 'i' have no contribution
- C_i
  - Default vector of constants
- Observations
  - In the worst case s_{-i} is empty
    - Agent 'i' impacts on all states
  - D_i directly expresses the impact of agent 'i' not present
  - Limited by capacity to design appropriate 'difference' expression

See also 'Factored Evolutionary Algorithms' (Strasser et al., 2017)

Cooperative Synapse NeuroEvolution

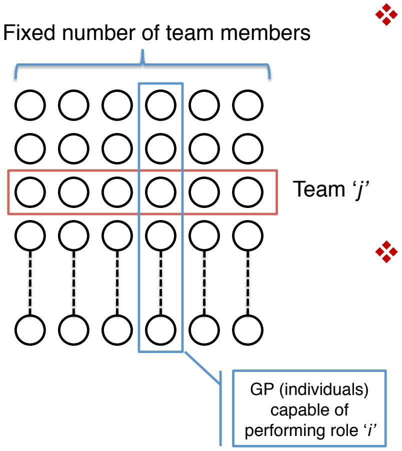
(Gomez et al., 2008)

- Select Parents
  - NNs (say, top 25%)
- Variation
  - 75% children
- Sort $\mathbf{P_i}$ w.r.t. $f(\boldsymbol{w_{ij}})$
  - $\mathbf{P_i} : f(\boldsymbol{w_{i1}}) > f(\boldsymbol{w_{i2}}) > \ldots f(\boldsymbol{w_{i\beta}})$
- Stochastic permutation of $\mathbf{P_i}$ content
  - $\mathbf{P_i} : f(\boldsymbol{w_{i1}})\ f(\boldsymbol{w_{i2}}) \ldots f(\boldsymbol{w_{i\beta}})$



Orthogonal evolution of (GP) teams (1)

(Thomason & Soule, 2007), (Rubini et al., 2009)



- **Motivation**
  - Team selection:
    - Good cooperation
    - Poor individual fitness
  - Island (individual) selection:
    - Poor cooperation
    - Strong individual fitness
- **OET1 (OET2)**
  - Select w.r.t individuals (teams)
  - Replace w.r.t. teams (individuals)

Orthogonal evolution of (GP) teams (2)

(Thomason & Soule, 2007), (Rubini et al., 2009)

OET1

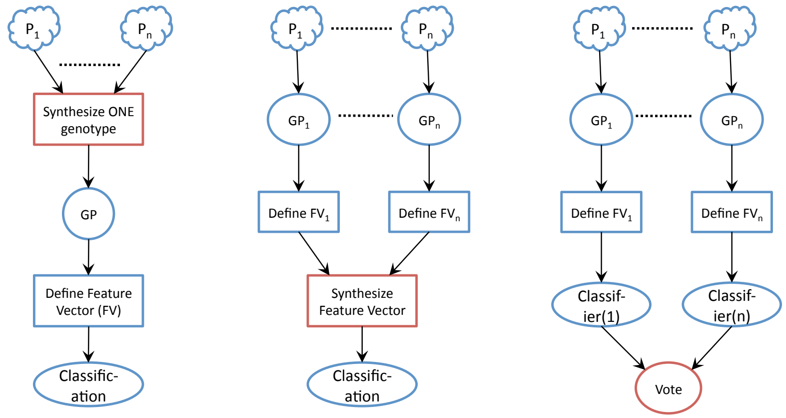
- Team = NULL
- *Select* **best** individual per role
- Create 2 such teams
- Apply variation operators
- Evaluate fitness
- *Replace* **worst teams**

OET2

- *Select* **2 best teams**
- Apply variation operators
- Evaluate fitness
- Award fitness to individuals in same team
- *Replace* **weakest** individuals

Level of Decomposition

(Krawiec & Bhanu, 2005), (Krawiec & Bhanu, 2007)



Agent Similarity/Diversity Ratio

(Gomes et al. 2018)

Coevolving a 5 agent team using 2 populations

Pop.1

Agent: A2
Policy: a_1, a_2, … a_{50}

Compose & evaluate a team

(A1, b7), (A2, a4), (A3, b7), (A4, b7)

Pop.2

Agent: A1, A3, A4
Policy: b_1, b_2, … b_{50}

Splitting (w.r.t. Pop. 1)

Pop.2a

Agent: A1
Policy: b_1, b_2, … b_{50}

Pop.2b

Agent: A3, A4
Policy: b_1, b_2, … b_{50}

Merge (w.r.t. Pop. 1 & 2a)

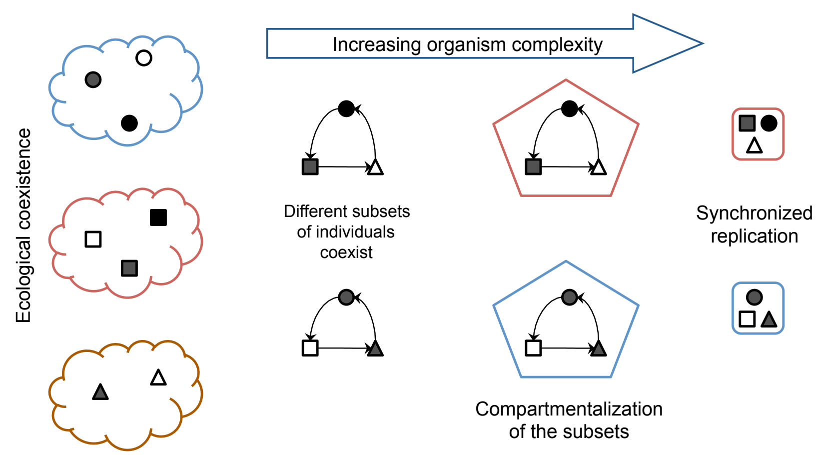
Agent: A1, A2
Policy: a_1, …, a_{30}, b_1 …, b_{20}

III.1 Case Study – Evolving arbitrary sized teams

Symbiosis, diversity maintenance, and separating action from context

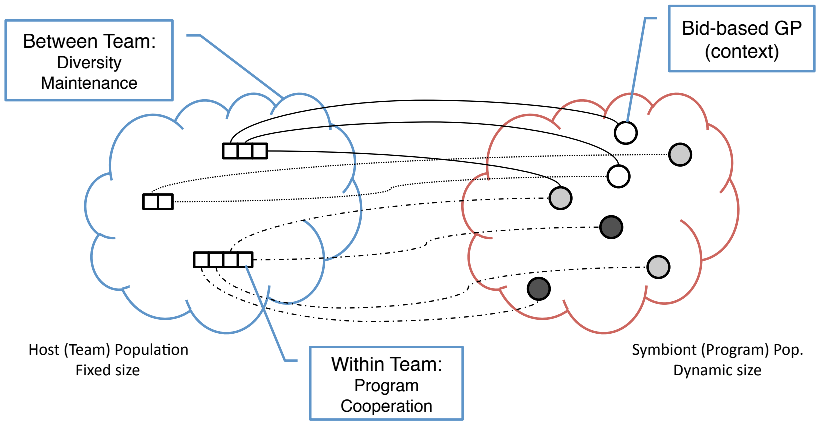
Abstract Model of Symbiosis

(Maynard Smith, 1991)



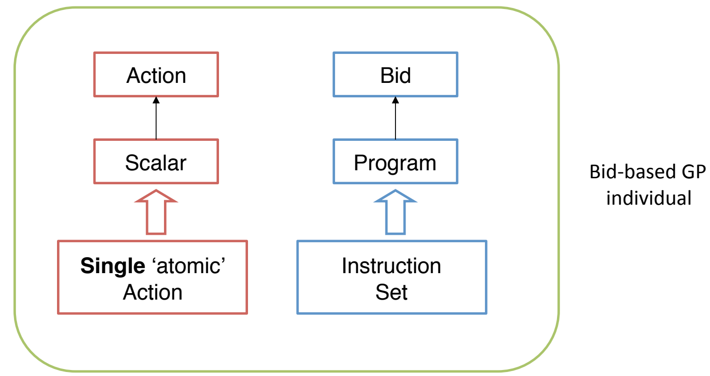
Symbiotic Bid-Based GP (SBB)

(Lichodzijewski & Heywood, 2008, 2010)



Achieving Symbiont Context

Bid-based GP



Team (Host) Fitness

- **Outcome vector, G(•)**
  - Point (p(k)) to Team/Host (h(i)) Outcome

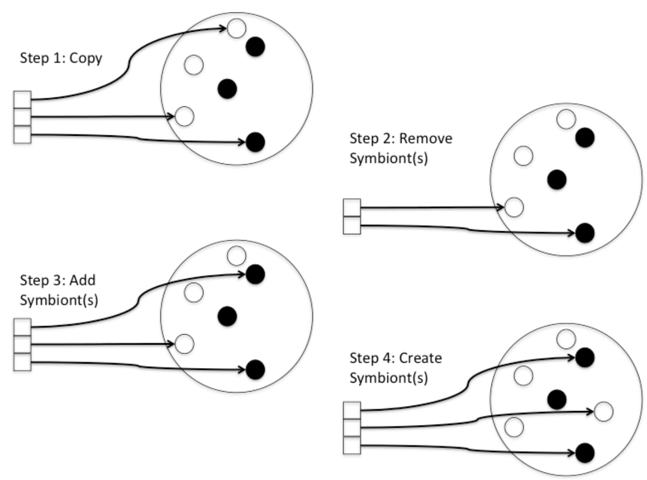
G(h(i), p(k)) = Domain specific reward on training case p(k)

- **Inter Host Diversity Maintenance**
  - Fitness sharing (see also behavioural and novelty measures)

$$s_i = \sum_k \left(\frac{G(h_i, p_k)}{\sum_j G(h_j, p_k)} \right)$$

Asexual Reproduction

Species independence



Supervised learning

- Multi-class classification
  - (Lichodzijewski & Heywood, 2008)
- Monolithic GP versus Teaming GP
  - (Lichodzijewski & Heywood, 2010)
- Decomposing large attribute spaces
  - (Doucette et al., 2012a)
- Streaming Classification
  - (Vahdat et al., 2015), (Khanchi et al., 2018)

III.2 Case Study – Non-stationary streams

Supporting Evolvability / Plasticity through Cooperative Coevolution

Non-stationary Streaming data

(Vahdat et al., 2015), (Khanchi et al., 2018)

Drift – 'gradual' variation

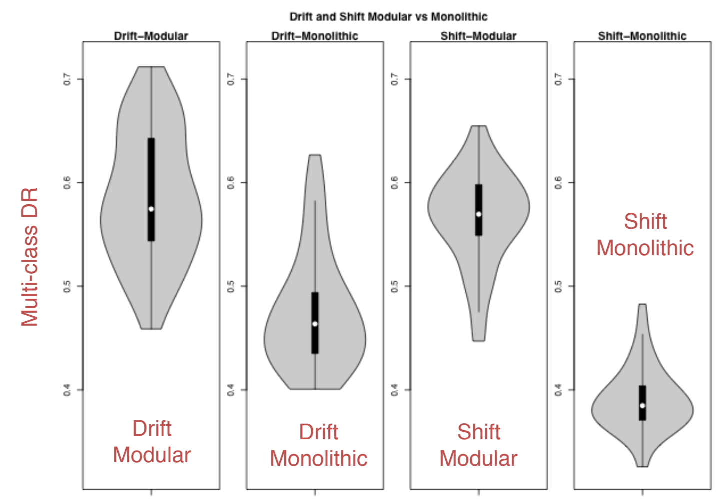
- 150,000 exemplars over stream
- Window interface
  - 500 window locations
  - 20 exemplars sampled per window location
- 10 attributes
- 3 classes
  - 16%, 74%, 10%

Shift – 'sudden' variation

- 6.5 million exemplars over stream
- Window interface
  - 1,000 window locations
  - 20 exemplars sampled per window location
- 6 attributes
- 5 classes
  - 36%, 49%, 6%, 0.5%, 1.5%, 3%, 4%

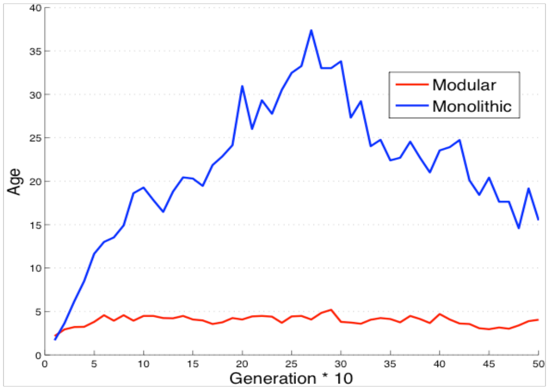
Accumulated multi-class detection rate

(Vahdat et al., 2015)



Age of champion individual

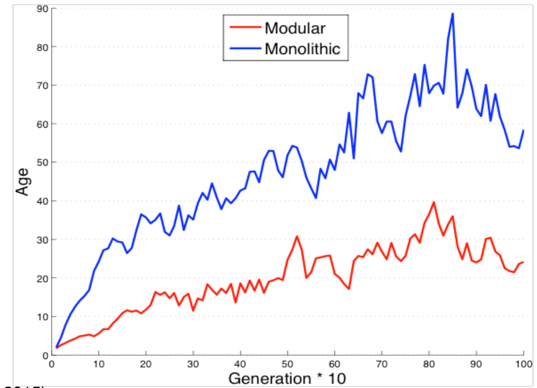
During course of stream – Drift



(Vahdat et al., 2015)

Age of champion individual

During course of stream – Shift



(Vahdat et al., 2015)

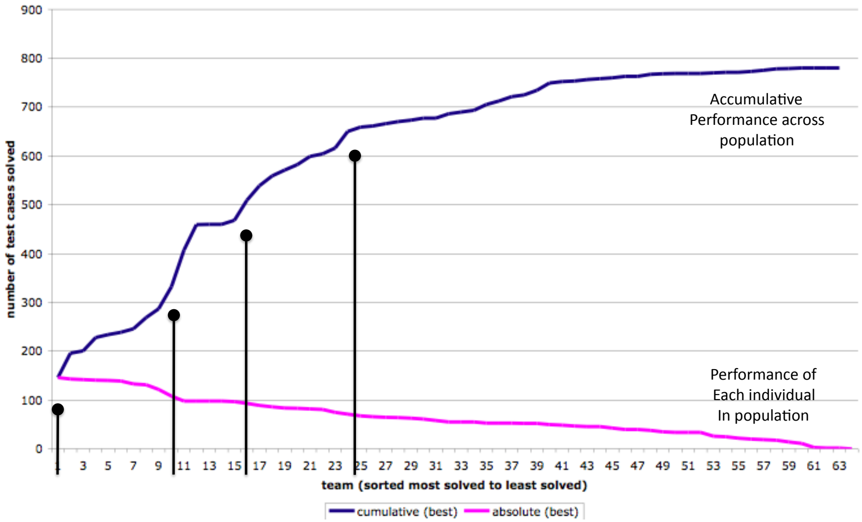
Observations

- **Context** for the symbiont **programs** must be evolved
  - Bidding mechanism
- Support for ***problem decomposition***
  - Mix of symbiont programs per host an evolved trait
    - No prior knowledge on the nature of an appropriate decomposition
    - Provides capacity for reacting to change
- Lower '**age**' of champion
  - Easier to switch in / out functional non-functional symbionts as contexts change
- Application: **Botnet detection** under label budgets
  - See (Khanchi et al., 2018)

III.3 Case Study – Diversity maintenance and Policy reuse

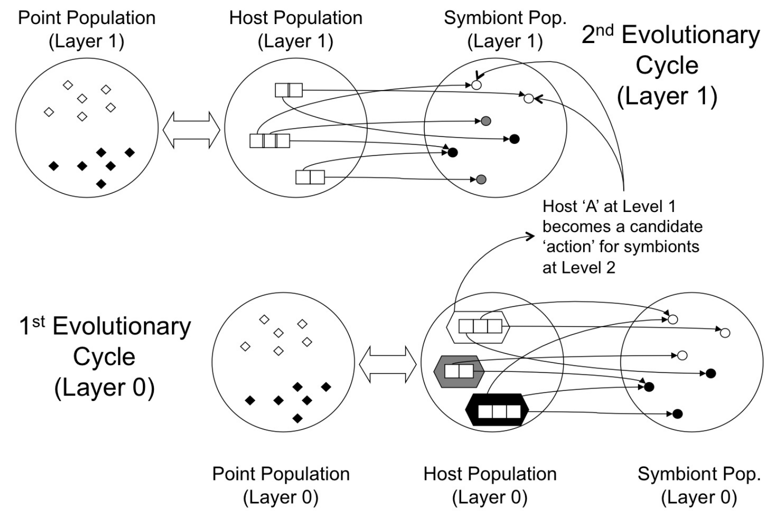
Hierarchical organization of programs, program abstraction

Motivation – Population fails in task



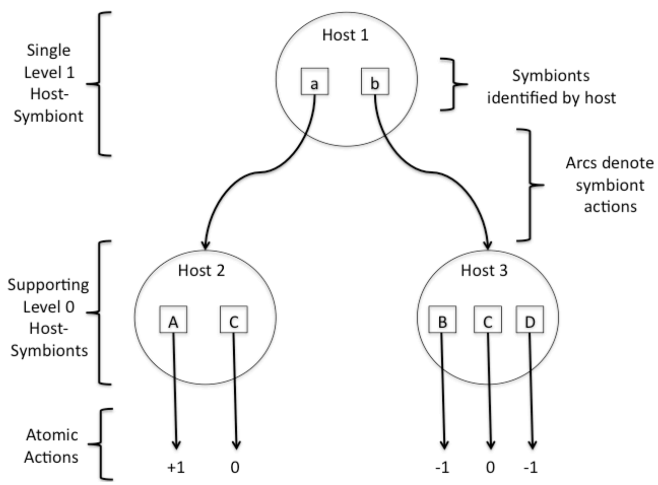
Evolving a policy tree

(Doucette et al., 2012b), (Kelly & Heywood 2014, 2015), (Smith et al, 2016)



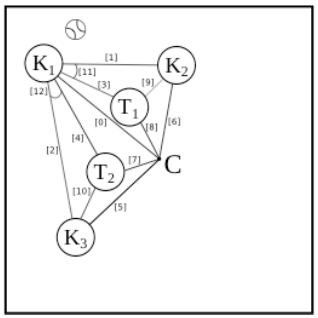
Evaluating a policy tree

(Doucette et al., 2012b), (Kelly & Heywood 2014, 2015, 2018a), (Smith et al, 2016)

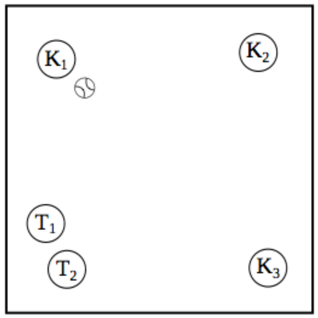


Keepaway soccer

Task definition (Stone et al, 2005)



Game initial state
-- Stochastically defined
-- Robocup server

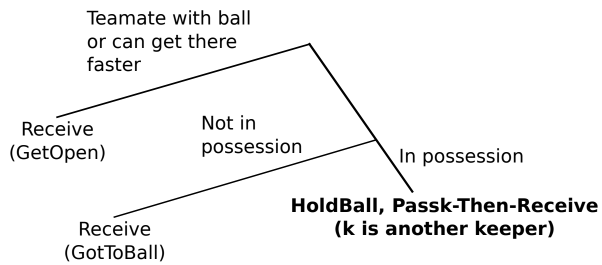


State variables
-- takers to keepers
-- ball assumes similar description

Interface to policy learner

Prior 'keeper' decision tree

Stone et al, (2005)



'Novelty' style diversity metric

(Kelly & Heywood, 2014)

- All start states the 'same'
- Encourage diversity in failure (novelty)

Reward of individual 'h_i' on game 'e_j'

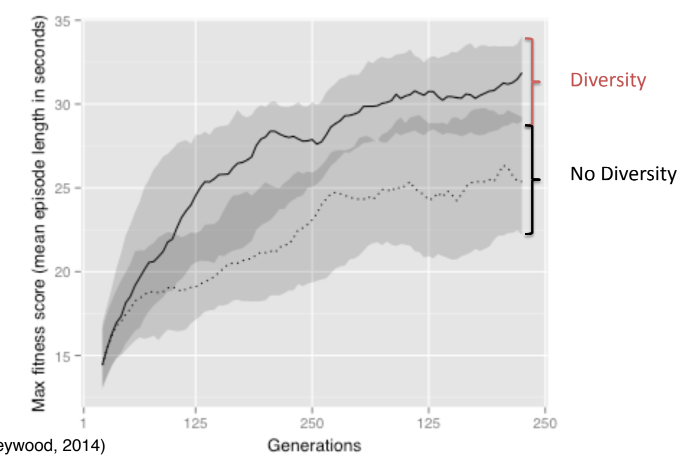
$$s_i = \sum_{j \in h_{hist}} \left(\frac{G(h_i, e_j)}{\sum_{k \neq i} (1 - dist(e_j, e_{hist})) G(h_k, e_{hist})} \right)$$

Distance between current game (e_j) and 'closest' historical game (ehist) for alternate solution (hk)

Reward of alternate individ. (h_k) in historical game (e_{hist})

Keepaway TRAINING performance

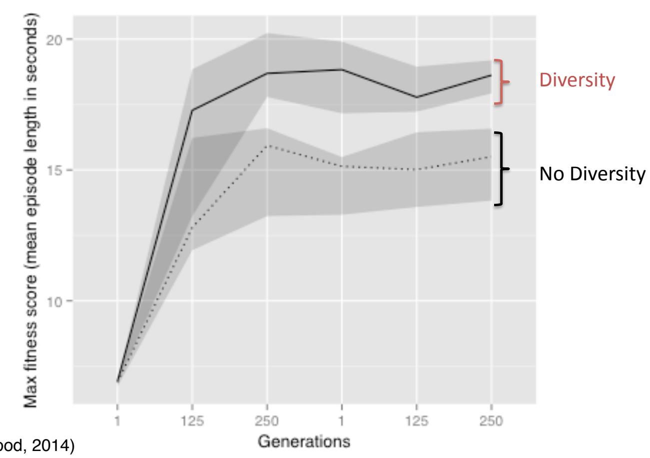
With / Without diversity



(Kelly & Heywood, 2014)

Keepaway TEST performance

1000 games, Sampled at intervals of 125 generations



(Kelly & Heywood, 2014)

III.4 Case Study – Multi-task learning under Arcade Learning Environment

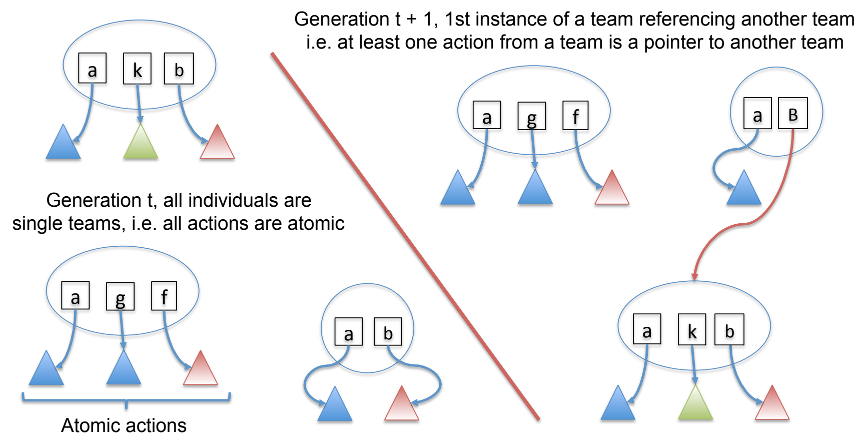
Program 'connectivity' organized through an emergent process

Tangled Program Graph representation

Single policies play multiple games

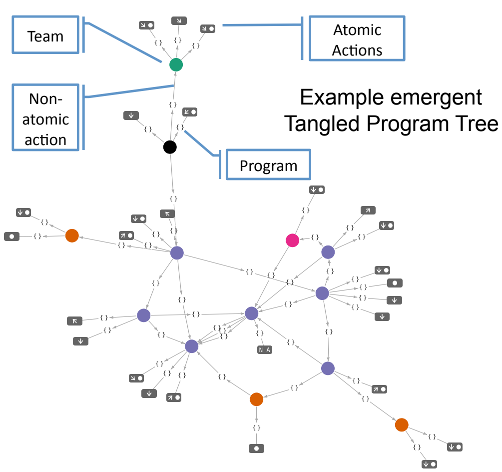
Tangled Program Graphs

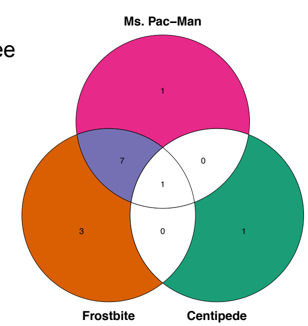
(Kelly and Heywood, 2017a, 2017b) (Smith and Heywood, 2018, 19a, 19b)



Playing *Multiple* Atari game titles

(Kelly & Heywood, 2017b)





Animation

(Kelly and Heywood, 2017a)

- Tutorial on emergent construction of Tangled Program Graphs
  - http://stephenkelly.ca/research_files/skelly-mheywood-eurogp-2017.pdf
- Solution TPG playing Frostbite title from Atari Learning Environment
  - http://stephenkelly.ca/research_files/TPG-frostbite-mosaic3.mp4

III.5 Case Study – Combining Competitive and Cooperative coevolution

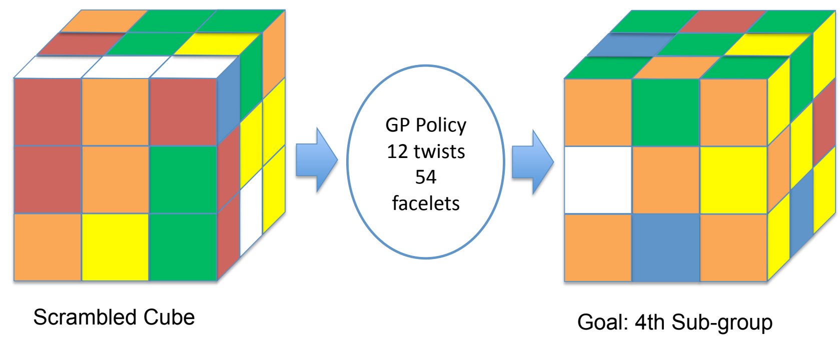
Evolving `Deep' policy hierarchies

Select Rubik's Cube configurations Competitively

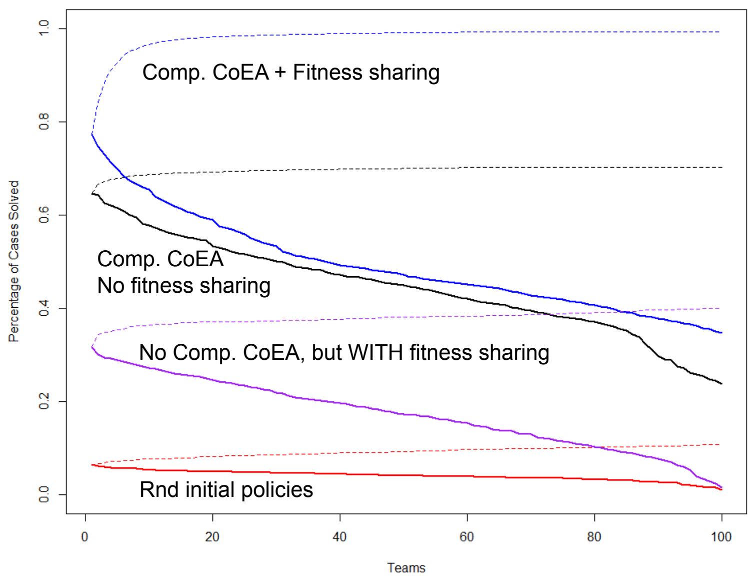
Coevolve teams of programs through independent cycles of evolution

Learning optimal policies for Rubik's Cube state

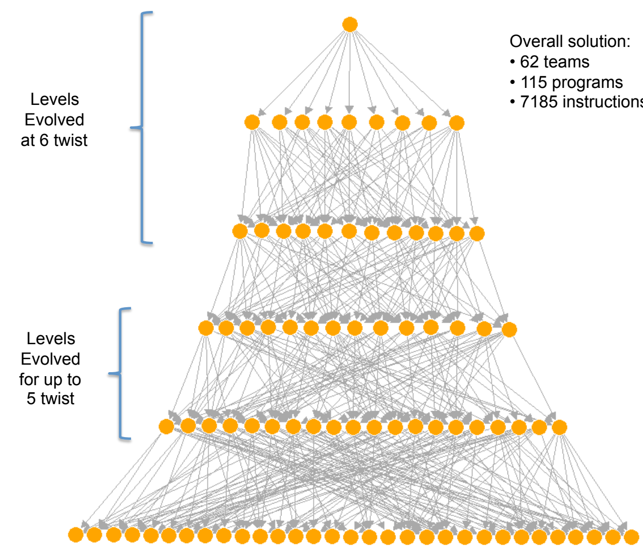
(Smith & Heywood, 2017)



Contribution of Competition and Diversity



Solution Complexity at 6-twist



Cooperative Coevolution
Concluding Comments (1 of 2)

- **Highlights**
  - Separation of context and action
    - Arbitrary team sizes under GP
  - Maintaining Diversity significant
    - Making diversity metrics 'task free'?
  - Reuse of previous policies leverages diversity for generalization
    - Strict cycles of reuse: **hierarchical policy trees**
    - Continuous discovery of modularity: **emergent tangled program graphs**
  - Solutions generally simpler than monolithic models
    - Real-time execution under modest computational support
  - React to changing environments more effectively

Cooperative Coevolution
Concluding Comments (1 of 2)

- **Some open questions (a non exhaustive list!)**
  - Credit for collective versus individuals
  - What learning bias are most appropriate for diversity maintenance
    - Task specific metrics
      - E.g., (Nelson et al. 2009)
    - … versus task independent metrics
      - Novelty as an objective (Gomes, Christensen 2013), (Gomes et al., 2016)
      - Compression distance (Gomez, 2009)
      - Connectivity biases (Clune et al., 2013)
      - Intra Team diversity (Kelly, Heywood, 2015), (Gomes et al., 2016)
    - … versus how to 'present' diversity
      - Pareto Multi-objective versus **switching** between multiple diversity metrics (Donieux, Mouret, 2013)
  - Cooperative coevolution and code reuse
    - Supervised learning (Lichodzijewski and Heywood, 2008, 2010), (Jaskowski et al., 2014)
    - Reinforcement learning (Kelly and Heywood, 2014, 2015, 2017a,b), (Didi and Nitschke, 2016), (Smith and Heywood, 2017, 2018)
  - Specialization versus generalization
    - Heterogeneous versus Homogeneous deployment of policies within teams (Waibel et al., 2009), (Nitschke et al., 2012)

Cooperative Coevolution
Example Benchmark task domains

- **Feature identification to classification**
  - K. Krawiec, B. Bhanu (2006, 2007); W. Jaskowski et al., (2014)
    - Constructing hierarchal models for feature extraction and classification
- **Double inverted pendulum / cart pole**
  - Gomez et al, (2008)
    - Capacity for solving the task
- **Acrobot**
  - Doucette et al, (2012b)
    - Capacity for solving the task / generalization
- **Predator-prey strategies**
  - Nitschke et al., (2012); Yong and Miikkulainen (2009); Rawael et al., (2010); Gomes et al., (2016)
    - Task decomposition and collective problem solving
- **Distributed multi-object location**
  - Agogino, Tumar (2008); Knudson, Tumar (2010); Colby, Tumar (2012)
    - Task decomposition and (heterogeneous) collective problem solving
- **Keepaway or Half field offense (soccer)**
  - Kelly, Heywood (2014, 2015, 2018a), Didi and Nitschke (2016), Gomes et al, (2018)
    - Task decomposition and collective problem solving
    - Capacity for task / generalization through hierarchical code reuse
- **Strategies for solving the Rubik's Cube**
  - Smith et al., (2016), Smith and Heywood (2017)
    - Task decomposition and capacity for task / generalization through hierarchical code reuse
- **General video game playing agents (i.e. Visual reinforcement learning)**
  - Arcade Learning Environment - Kelly and Heywood (2017a,17b, 18b))
    - Comparison against Deep Learning Visual RL agents
  - VizDoom FPS - Smith and Heywood (2019a) and Dota 2 Shadow Fiend Bot - Smith and Heywood (2019b)
    - Synchronizing external memory for addressing partially observable state

IV. Closing remarks

Closing remarks

- Coevolutionary algorithms = conceptually interesting and oftentimes efficient paradigm for solving complex problems
- Addresses key aspects of computational intelligence:
  - What/who to learn from?
  - How to drive the search/optimization?
  - What is solution to my problem?
  - How do I decompose my problem?
  - How do I make some entities cooperate?
- Many interesting results,
  - … even more open questions!

Acknowledgements

- The content of this tutorial has benefited from a host of collaborations over the years including, but not limited to:

  John Doucette, Wojciech Jaśkowski, Stephen Kelly, Peter Lichodzijewski, Paweł Liskowski, Robert Smith, Marcin Szubert, Ali Vahdat, Bartosz Wieloch
- MIH would like to acknowledge funding for aspects of research reported on in this tutorial from the NSERC Discovery and CRD programs (Canada).
- KK would like to acknowledge funding for aspects of research reported on in this tutorial from the National Science Centre and National Centre for Research and Development (Poland, grant 2014/15/B/ST6/05205).

References

Competitive Coevolution (1 of 3)

- A. Arcuri, X. Yao, *Co-evolutionary automatic programming for software development*, Information Sciences, 259:412-432, February 2014.
- R. Axelrod (1987) *The evolution of strategies in the iterated prisoner's dilemma.* In L. Davis, editor, Genetic Algorithms in Simulated Annealing, 32–41. Pitman, London.
- W.W. Bledsoe, I. Browning (1959) Pattern recognition and reading by machine. In Proc. Eastern Joint Comput. Conf., 225–232.
- A. Bucci, J.B. Pollack, E. de Jong (2004) *Automated extraction of problem structure.* In K. Deb et al. (Eds.), Genetic and Evolutionary Computation, GECCO-2004, Part I. Lecture Notes in Computer Science, Vol. 3102, 501–512. Berlin: Springer-Verlag
- S. Y. Chong, P. Tino, and X. Yao (2008) *Measuring generalization performance in coevolutionary learning,* IEEE Trans. Evol. Comput., vol. 12, no. 4, 479–505
- S.G. Ficici (2004) *Solution concepts in coevolutionary algorithms*, Ph.D. thesis, Brandeis University, Waltham, MA.
- S.G. Ficici, J.B. Pollack (2001) *Pareto optimality in coevolutionary learning.* In J. Kelemen and P. Sosık (Eds.), Advances in Artificial Life, 6th European Conference, ECAL'01. Lecture Notes in Computer Science, Vol. 2159, 316–325. Berlin: Springer-Verlag
- D.B. Fogel (2002) *Blondie24: Playing at the Edge of AI,* Morgan Kaufmann Publishers Inc., San Francisco, CA.
- W. Jaśkowski, K. Krawiec and B. Wieloch (2008) *Evolving Strategy for a Probabilistic Game of Imperfect Information using Genetic Programming.* Genetic Programming and Evolvable Machines, 9(4):281-294
- W. Jaśkowski, K. Krawiec (2010) *Coordinate System Archive for coevolution.* In IEEE Congress on Evolutionary Computation.
- W. Jaśkowski, K. Krawiec (2011) *How many dimensions in co-optimization.* In GECCO (Companion), 829-830.
- W. Jaśkowski, K.Krawiec (2011) *Formal Analysis, Hardness, and Algorithms for Extracting Internal Structure of Test-Based Problems.* Evolutionary Computation, 19(4):639-671.
- E.D. de Jong (2004) *Towards a Bounded Pareto-Coevolution Archive.* In Proceedings of the IEEE Congress on Evolutionary Computation, volume 2, 2341– 2348, Portland, Oregon, USA.
- E.D. de Jong (2004) *The Incremental Pareto-Coevolution Archive.* In GECCO 2004. Proceedings of the Genetic and Evolutionary Computation Conference. Part I, 525–536, Springer-Verlag, Lecture Notes in Computer Science Vol. 3102.
- E.D. de Jong, J.B. Pollack (2004) *Ideal evaluation from coevolution.* Evolutionary Computation, 12(2):159–192.
- E. D. de Jong (2004) *The MaxSolve algorithm for coevolution,* in GECCO 2005: Proceedings of the 2005 conference on Genetic and evolutionary computation, 2005, 483–489.

References

Competitive Coevolution (2 of 3)

- E.D. de Jong, A. Bucci (2006) *DECA: Dimension extracting coevolutionary algorithm.* In Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, GECCO 2006, 313–320
- K. Krawiec, (2001) *Pairwise Comparison of Hypotheses in Evolutionary Learning.* In Machine Learning. Proceedings of the Eighteenth International Conference, ICML 2001. Morgan Kaufmann Publishers, 266-273.
- K. Krawiec, I. Bladek, J. Swan (2017) *Counterexample-Driven Genetic Programming*, ACM GECCO'17. 953-960
- K. Krawiec, P. Liskowski (2015) *Automatic Derivation of Search Objectives for Test-Based Genetic Programming,* in P. Machado, M. Heywood, J. McDermott (eds.), 18th European Conference on Genetic Programming, Springer
- K. Krawiec and M. Szubert (2011) *Learning N-tuple networks for Othello by coevolutionary gradient search,* in Proc. Genetic Evol. Comput. Conf., ACM 355–362.
- P. Liskowski, K. Krawiec (2016). *Non-negative Matrix Factorization for Unsupervised Derivation of Search Objectives in Genetic Programming.* ACM GECCO'16, 749-756
- P. Liskowski, K. Krawiec (2016), *Surrogate Fitness via Factorization of Interaction Matrix.* EuroGP'16, LNCS, Springer. 68-82
- P. Liskowski, K. Krawiec (2017) *Online Discovery of Search Objectives for Test-based Problems.* Evolutionary Computation Journal, MIT Press, 25(3): 375-406.
- P. Liskowski, K. Krawiec, B. Wieloch (2018), Neural Estimation of Interaction Outcomes, ACM GECCO'18.
- T. Miconi (2009) *Why coevolution doesn't work: Superiority and progress in coevolution,* In: L. Vanneschi, et al. (eds.), EuroGP 2009, Springer-Verlag, Berlin Heidelberg New York, 49–60.
- G.A. Monroy, K.O. Stanley, and R. Miikkulainen (2006) *Coevolution of neural networks using a layered Pareto archive.* In M. Keijzer et al., editors, GECCO 2006: Proceedings of the 8th annual conference on Genetic and evolutionary computation, volume 1, 329–336, Seattle, Washington, USA, 8-12 July 2006. ACM Press.
- P. Moscato (1989) *On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms,* Caltech Concurrent Computation Program C3P Rep., vol. 826.
- J. Noble, R.A. Watson (2001) *Pareto coevolution: Using performance against coevolved opponents in a game as dimensions for Pareto selection.* In L. Spector et al. (Eds.), Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2001, 493–500.

References
Competitive Coevolution (3 of 3)

- J.B. Pollack, A.D. Blair (1998) *Co-evolution in the successful learning of backgammon strategy.* Mach. Learn. 32(3), 225–240
- E. Popovici, A. Bucci, R.P. Wiegand, and E.D. de Jong (2012) *Coevolutionary Principles.* In Rozenberg, G., Baeck, T., and Kok, J. N., editors, Handbook of Natural Computing, 987–1033. Springer.
- E. Popovici, E. Winston (2015) *A framework for co-optimization algorithm performance and its application to worst-case optimization,* Theoretical Computer Science, Volume 567, Pages 46-73
- E. Popovici, *Bridging Supervised Learning and Test-Based Co-optimization*, JMLR, 18(38):1-39.
- C.D. Rosin and R. K. Belew (1997) *New methods for competitive coevolution,* Evolutionary Computation, vol. 5, no. 1, 1–29.
- A.L. Samuel (1959) *Some studies in machine learning using the game of checkers.* IBM Journal of Research and Development, 3(3):211–229.
- O.G. Selfridge, R.S. Sutton, A.G. Barto (1985) *Training and Tracking in Robotics.* In Joshi, A. K., editor, Proceedings of the 9th International Joint Conference on Artificial Intelligence, IJCAI, 670–672, Los Angeles, CA. Morgan Kaufmann.
- T.C. Service, D.R. Tauritz (2008) *A no-free-lunch framework for coevolution,* in: Proceedings of the Genetic and Evolutionary Computation Conference, ACM, 371–378.
- M. Szubert, Coevolutionary (2014) *Shaping for Reinforcement Learning,* Phd Thesis, Institute of Computing Science, Poznan University of Technology.
- M. Szubert, W. Jaśkowski, K. Krawiec (2009) *Coevolutionary Temporal Difference Learning for Othello.* In IEEE Symposium on Computational Intelligence and Games. 104-111.
- M. Szubert, W. Jaśkowski, P. Liskowski, K. Krawiec (2013) *Shaping Fitness Function for Evolutionary Learning of Game Strategies.* In Proceeding of the Fifteenth Annual Conference on Genetic and Evolutionary Computation Conference, GECCO '13, 1149–1156, New York, NY, USA. ACM.
- M. Szubert, W. Jaśkowski, K. Krawiec (2013) *On Scalability, Generalization, and Hybridization of Coevolutionary Learning: A Case Study for Othello.* Computational Intelligence and AI in Games, IEEE Transactions on, 5(3):214-226.
- B. F. Skinner (1938) *The behavior of organisms: An experimental analysis.* Appleton-Century.
- D. Wolpert, W. Macready (2005) *Coevolutionary free lunches,* IEEE Trans. Evol. Comput. 9: 721–735.

References
Cooperative Coevolution (1 of 3)

- A. K. Agogino, K. Tumar (2008) *Efficient evaluation functions for evolving coordination.* Evolutionary Computation 16(2): 257–288
- J. Clune, J.-B. Mouret, H. Lipson (2013) *The evolutionary origins of modularity.* Proceedings of the Royal Society – B 280 20122863
- S. Didi, G. Nitschke (2016) *Multi-agent behavior based policy transfer.* EvoApplications. LNCS 9598: 181—197
- M. Colby, K. Tumer (2012) *Shaping fitness functions for coevolving cooperative multiagent systems.* ACM AAMAS 425–432
- S. Doncieux, J.-B. Mouret (2013) *Behavioral diversity with multiple behavioral distances.* IEEE CEC 1–8
- J. Gomes, P. Mariano, A. L. Christensen (2017) *Novelty-driven cooperative coevolution.* Evolutionary Computation. 25(2): 275-307
- J. Gomes, A. L. Christensen (2014) *Generic behaviour similarity measures for evolutionary swarm robotics.* ACM GECCO 199—206
- J. Gomes, P. Mariano, A. L. Christensen (2018) Dynamic team heterogeneity in cooperative coevolutionary algorithms. IEEE Transaction on Evolutionary Computation. 22(6): 934-948
- F. Gomez, J. Schmidhuber, R. Miikkulainen (2008) *Accelerated neural evolution through cooperatively coevolved synapses.* Journal of Machine Learning Research 9:937–965
- F. Gomez (2009) *Sustaining diversity using behavioural information distance.* ACM GECCO 113–120
- S. Luke, C. Hohn, J. Farris, G. Jackson, J. Hendler (1997) Coevolving soccer softbot team coordination with genetic programming. IJCAI. RoboCup Workshop: 214-222
- W. Jaskowski, K. Krawiec, B. Wieloch (2014) *Cross-task code reuse in genetic programming applied to visual learning.* Applied Mathematics and Computer Science 24(1): 183—197
- M. Knudson, K. Tumar (2010) *Coevolution of heterogeneous multi-robot teams.* ACM GECCO 127–132
- K. Krawiec, B. Bhanu (2007) *Visual learning by evolutionary and coevolutionary feature synthesis.* IEEE Transactions on Evolutionary Computation 11(5): 635–650
- K. Krawiec, B. Bhanu (2006) *Visual learning by coevolutionary feature synthesis.* IEEE Transactions on Systems, Man and Cybernetics. Prt B. 35: 409–425
- J. Maynard Smith (1991) *A Darwinian view of symbiosis.* Chapter 3 in Symbiosis as a source of evolutionary innovation. (eds) L. Margulis and R. Fester (MIT Press)

References
Cooperative Coevolution (2 of 3)

- D. E. Moriarty, R. Miikkulainen (1998) *Forming neural networks through efficient and adaptive coevolution.* Evolutionary Computation 5(4):373–399
- A. L. Nelson, G. J. Barlow, L. Doitsidis (2009) *Fitness functions in evolutionary robotics: A survey and analysis.* Robotics and Autonomous Systems 57: 345–370
- G. S. Nitschke, A. E. Eiben, M. C. Schut (2012) *Evolving team behaviors with specialization.* Genetic Programming and Evolvable Machines 13(4): 493—536
- L. Panait, S. Luke, R. P. Wiegand (2006) *Biasing coevolutionary search for optimal multiagent behaviors.* IEEE Transactions on Evolutionary Computation 10(6): 629–645
- L. Panait, K. Tuyls, S. Luke (2008) *Theoretical advantages of lenient learners: An evolutionary game theoretic perspective.* Journal of Machine Learning Research 9: 423–457
- M. A. Potter, K. A. De Jong (2000) *Cooperative coevolution: An architecture for coevolving coadapted subcomponents.* Evolutionary Computation 8(1): 1–29
- A. Rawal, P. Rajagoplan, R. Miikkulainen (2010) *Constructing competitive and cooperative agent behavior using coevolution.* IEEE CIG 107—114
- J. Rubini, R. B. Heckendorn, T. Soule (2009) *Evolution of team composition in multi-agent systems.* ACM GECCO 1067–1072
- P. Stone, R. Sutton, G. Kuhlmann (2005) *Reinforcement learning for RoboCup soccer Keepaway.* Adaptive Behavior 13: 165–188
- S. Strasser, J. Sheppard, N. Fortier, R. Goodman (2017) *Factored Evolutionary Algorithms.* IEEE Transactions on Evolutionary Computation. 21(2): 281-293
- R. Thomason, T. Soule (2007) *Novel ways of improving cooperation and performance in ensemble classifiers.* ACM GECCO 1708–1716
- C. H. Yong and R. Miikkulainen (2009) *Coevolution of role-based cooperation in multi-agent systems.* IEEE Transactions on Autonomous Mental Development 1(3): 170—186
- M. Waibel, L. Keller, D. Floreano (2009) *Genetic team composition and level of selection in the evolution of cooperation.* IEEE Transactions on Evolutionary Computation. 13(3):648—660
- S. Wu, W. Banzhaf (2011) *Rethinking multilevel selection in genetic programming.* ACM GECCO. 1403 – 1410

References
Cooperative Coevolution (3 of 3)

- J.A. Doucette, A. R. McIntyre, P. Lichodzijewski, M. I. Heywood (2012a) *Symbiotic coevolutionary genetic programming: A benchmarking study under large attribute spaces.* Genetic programming and Evolvable Machines. 13(1): 71-101
- J. A. Doucette, P. Lichodzijewski, M. I. Heywood (2012b) *Hierarchical task decomposition through symbiosis in reinforcement learning.* ACM GECCO 97–104
- S. Khanchi, A. Vahdat, M. I. Heywood, A. N. Zincir-Heywood (2018) *On Botnet detection with GP under streaming data, label budgets and class imbalance.* Swarm and Evolutionary Computation. 39:123-140
- S. Kelly, M.I. Heywood (2014) *On diversity, teaming, and hierarchical policies: Observations from the Keepaway soccer task.* EuroGP LNCS 8599:75–86
- S. Kelly, M.I. Heywood (2015) *Knowledge transfer from keepaway soccer to half- field offense through program symbiosis: Building simple programs for a complex task.* ACM GECCO. 1143-1150
- S. Kelly, M.I. Heywood (2017a) *Emergent Tangled Graph Representations for Atari game playing agents.* EuroGP. LNCS 10196: 64-79
- S. Kelly, M.I. Heywood (2017b) *Multi-task learning in Atari video games with Emergent Tangled Program Graphs.* ACM GECCO. 195-202
- S. Kelly, M.I. Heywood (2018a) *Discovering agent behaviours through code reuse: Examples from half field offense and Ms. Pac-Man.* IEEE Transactions on Games 10(2): 195-208
- S. Kelly, M.I. Heywood (2018b) Emergent solutions to high-dimensional multitask reinforcement learning. Evolutionary Computation 26(3): 347-380
- Lichodzijewski, M. I. Heywood (2008) *Managing team-based problem solving with symbiotic bid-based genetic programming.* ACM GECCO 363–370
- P. Lichodzijewski, M. I. Heywood (2010) *Symbiosis, Complexification and Simplicity under GP.* ACM GECCO 853–860
- R. J. Smith, S. Kelly, M. I. Heywood (2016) *Discovering Rubik's Cube Subgroups using Coevolutionary GP -- A Five Twist Experiment.* ACM GECCO. 789-796
- R. J. Smith, M. I. Heywood (2017) *Coevolving deep hierarchies of programs to solve complex tasks.* ACM GECCO. 1009-1016
- R. J. Smith, M. I. Heywood (2018) *Scaling tangled program graphs to visual reinforcement learning in VizDoom.* EuroGP. LNCS 10781: 135-150
- R. J. Smith, M. I. Heywood (2019a) Model of external memory for navigation in partially observable visual reinforcement learning tasks. EuroGP. LNCS 11451
- R. J. Smith, M. I. Heywood (2019b) Evolving Dota 2 shadow fiend bots using genetic programming with external memory. ACM GECCO.
- A. Vahdat, J. Miller, A. McIntyre, M. I. Heywood, N. Zincir-Heywood (2015) *Evolving GP classifiers for streaming data tasks with concept change and label budgets.* Handbook of GP Applications. (Springer)