COMP 5660/6660 Fall 2025 Final Exam Key

This is a closed-book, closed-notes exam. The sum of the max points for all the questions is 130, but note that the max exam score will be capped at 122 (i.e., there are 8 bonus points, but you can't score more than 100%). You have exactly two hours to complete this exam. Good luck!

Multiple Choice Questions

1. While in a standard EA an offspring is generated by recombination followed by mutation, in GP one usually generates an offspring either by recombination or by mutating a clone of a parent, not both. This is because: [4 pts]

   (a) the combination of recombination and mutation frequently creates too much stochastic noise, effectively resulting in random search; GP is a relatively new type of EA which allowed its creators to correct this problem by designing it from the start to do either recombination or mutation, but not both at the same time

   (b) recombination and mutation are often quite destructive in GP and doing both would effectively result in random search

   (c) performing both recombination and mutation would violate the closure property of GP

   Select one of:

   - a [2]
   - **b**
   - c [0]
   - a and b [3]
   - a and c [1]
   - b and c [2]
   - a, b, and c [3]
   - none of a, b, nor c [0]

2. Which of the following is an example of test-based competitive coevolution? [4 pts]

   (a) coevolving Pac-Man controllers and maps

   (b) coevolving Pac-Man controllers and ghost controllers **(example of interactive competitive coevolution)**

   (c) coevolving ghost controllers and maps **(example of cooperative coevolution)**

   Select one of:

   - **a**
   - b [2]
   - c [1]
   - a and b [3]
   - a and c [2]
   - b and c [1]
   - all of a, b, and c [2]
   - none of a, b, nor c [0]

3. In the context of Assignment 2c, the risk of evaluating a Pac-Man controller against a single Ghost controller is: [4 pts]

   (a) As this undersamples the opposing population in the most extreme, the Pac-Man controller's fitness estimation is almost guaranteed not representative of the true value.

   (b) There may be another Ghost who is better than the Ghost sampled against, yet Pac-Man does better against that Ghost, violating the transitive property of the Ghost fitness function.

   (c) While it does undersample the opposing population, the greatly reduced evaluation time allows for many more generations to be run, and this is always a beneficial tradeoff.

   Select one of:

   - **a**
   - b [1]
   - c [2]
   - a and b [2]
   - a and c [3]
   - b and c [1]
   - a, b, and c [2]
   - none of a, b, nor c [0]

4. Modern Evolutionary Programming (EP) differs from classic EP in: [4 pts]

   (a) representation

   (b) parent selection

   (c) parameter control

   Select one of:

   - a [2]
   - b [0]
   - c [2]
   - a and b [1]
   - **a and c**
   - b and c [1]
   - a, b, and c [3]
   - none of a, b, nor c [0]

5. Koza states that a parameterized topology in GP is a: [4 pts]

   (a) general solution to a problem in the form of a graphical structure whose nodes or edges represent components and where the parameter values of the components are specified by mathematical expressions containing free variables

   (b) search landscape for tree representations whose terminal nodes take the values of input parameters

   (c) graph representation where terminal node input values are determined employing parameter control

   Select one of:

   - **a**
   - b [2]
   - c [0]
   - none of a, b, nor c [0]

6. Panmictic mate selection in EAs has the following properties: [4 pts]

   (a) strategy parameters are fixed during an EA run
   (b) no genotypic restrictions on mating
   (c) more fit individuals mate more often
   (d) process of tuning mate selection parameters for each problem is time-consuming

   Select one of:

   - a [0]
   - **b**
   - c [0]
   - d [0]
   - a and b [1]
   - a and c [0]
   - a and d [0]
   - b and c [2]
   - b and d [2]
   - c and d [0]
   - a, b, and c [1]
   - a, b, and d [1]
   - a, c, and d [0]
   - b, c, and d [1]
   - all of a, b, c, and d [1]
   - none of a, b, c, nor d [0]

7. In Crowding: [4 pts]

   (a) new individuals replace similar members of the population, resulting in the population being evenly distributed over the niches regardless of niche fitness
   (b) the fitness of individuals immediately prior to selection is adjusted according to the number of individuals falling within some prespecified distance of each other
   (c) individuals share the fitness of similar population members immediately prior to selection, resulting in the number of individuals per niche being dependent on the niche fitness

   Select one of:

   - **a**
   - b [1]
   - c [2]
   - a and b [2]
   - a and c [3]
   - b and c [1]
   - all of a, b, and c [2]
   - none of a, b, nor c [0]

8. Fitness sharing differs from crowding in that fitness sharing: [4]

    (a) results in panmictic mating
    (b) results in niches sized proportional to fitness
    (c) implicitly requires fitness proportionate selection
    (d) implicitly requires fitness ranked selection

    Select one of:

    - a [0]
    - b [2]
    - c [2]
    - d [0]
    - a and b [1]
    - a and c [1]
    - a and d [0]
    - **b and c**
    - b and d [1]
    - c and d [1]
    - none of a, b, c, nor d [0]

9. Speciation is: [4 pts]

    (a) when geographically separated sub-populations of a species adapt to their local environmental niches to the extent that they become mating-incompatible
    (b) when geographically separated sub-populations of a species adapt to their local environmental niches to the extent that they become mating-compatible
    (c) when sub-populations of different species in the same local environmental niche adapt homogeneously to the extent that they become mating-incompatible
    (d) when sub-populations of different species in the same local environmental niche adapt homogeneously to the extent that they become mating-compatible

    Select one of:

    - **a**
    - b [2]
    - c [0]
    - d [1]
    - a and c [1]
    - a and d [2]
    - b and c [0]
    - b and d [1]
    - none of a, b, c, nor d [0]

10. According to the concept of island model EAs in the context of Eldredge and Gould's theory of punctuated equilibria: [4 pts]

    (a) multiple populations of different species are run in parallel in some kind of communication structure

    (b) after a usually variable number of generations, a number of individuals are selected from each population to be exchanged with others from neighboring populations

    (c) during the migration phase, the injection of individuals of potentially high fitness, and with possibly radically different genotypes, facilitates exploration

    (d) the migratory injections interrupt periods of evolutionary stasis by rapid growth when the main population is invaded by individuals from a previously spatially isolated group of individuals of a different species.

    Select one of:

    - a [2]
    - b [2]
    - **c**
    - d [3]
    - all of a, b, c, and d [2]
    - none of a, b, c, nor d [0]

11. In competitive coevolution, the Hall of Fame is meant to address which of the following stability pathologies: [4 pts]

    (a) cycling

    (b) mediocre stability

    (c) disengagement

    (d) overspecialization

    Select one of:

    - **a**
    - b [0]
    - c [0]
    - d [0]
    - a and b [2]
    - a and c [2]
    - a and d [2]
    - b and c [0]
    - b and d [0]
    - c and d [0]
    - a, b, and c [1]
    - a, b, and d [1]
    - a, c, and d [1]
    - b, c, and d [0]
    - all of a, b, c, and d [0]
    - none of a, b, c, nor d [0]

12. Symmetric and asymmetric games can differ in: [4 pts]

    (a) commonality of player objectives
    (b) commonality of player capabilities
    (c) number of populations needed by competitive coevolution

    Select one of:

    - a [1]
    - b [1]
    - c [1]
    - a and b [3]
    - a and c [3]
    - b and c [3]
    - **a, b, and c**
    - none of a, b, nor c [0]

13. The challenge when employing parameter control in order to reduce the number of EA strategy parameters which the practitioner has to configure is: [4 pts]

    (a) the introduction of "stealth" parameters, namely new parameters to control the parameter control, which may be as hard or harder to tune than the parameter(s) eliminated by the employment of the parameter control
    (b) the introduction of "stealth" parameters, namely new parameters to control the parameter control, which cause a dynamic derived variable associated with the eliminated EA strategy parameter to converge to a fixed value deterministically specified by the stealth parameters
    (c) the interaction between the parameter control of different operators, such as population sizing and offspring sizing, may be complex and hard to tune

    Select one of:

    - a [2]
    - b [2]
    - c [2]
    - a and b [3]
    - a and c [3]
    - b and c [3]
    - **all of a, b, and c**
    - none of a, b, nor c [0]

Open Questions

14. Explain why it does or why it does not make sense to investigate the Baldwin Effect for a Lamarckian Evolutionary Algorithm approach to solving the Assignment Series 1 Cutting Stock Problem. [2 pts]

    This does not make sense, because Lamarckian EAs do not exhibit the Baldwin Effect.

15. How are conflicting rules in the action set of a Learning Classifier System resolved? [2 pts]

    This is not applicable because per definition all the rules in the action set advocate the same action.

16. Is the genotypic encoding for the Assignment 2 Series of Pac-Man vs. the Ghosts pleiotropic, polygenic, both, or neither? Explain your answer! [6 pts]

    It is pleiotropic and polygenic, because one gene (function or terminal node in GP tree) can impact multiple phenotypic traits (controller actions in the form of GP tree outputs) which means the genotypic encoding is pleiotropic, and one phenotypic trait (controller action) can depend on multiple genes (function or terminal nodes in GP tree) which means the genotypic encoding is polygenic.

17. Is the genotype-phenotype decoding function for the Assignment 2 Series of Pac-Man vs. the Ghosts surjective, injective, both, or neither? Explain your answer! [6 pts]

    It is surjective but not injective, because all controllers are valid genotypes (surjective), but there exist controllers than can be encoded by multiple distinct genotypes, for instance by swapping two constant terminals being fed into a summation function (not injective).

18. Is the phenotype to fitness mapping for the Assignment 2 Series of Pac-Man vs. the Ghosts surjective, injective, both, or neither? Explain your answer! [6 pts]

    It is surjective but not injective, because potentially all controllers can be represented, and therefore all valid fitness values obtained (surjective), but there exist distinct controllers which obtain the same fitness, for instance given a symmetric scenario they follow a reverse direction strategy (not injective).

19. Assuming an elitist evolutionary algorithm with whose global optimum has a fitness of 100 and given a population consisting of the following bit strings v_1 through v_5 and schema S
    $v_1 = (11110110011001)$ $fitness(v_1) = 19$
    $v_2 = (10110111011011)$ $fitness(v_2) = 11$
    $v_3 = (10110110011101)$ $fitness(v_3) = 2$
    $v_4 = (10111110011001)$ $fitness(v_4) = 100$
    $v_5 = (11111110111001)$ $fitness(v_5) = 10$
    $S = (10111110011001)$

    (a) Compute the *order* of S. [1] *14*
    (b) Compute the *defining length* of S and show your computation. [1] *14-1=13*
    (c) Compute the fitness of S and show your computation. [1] $fitness(S) = fitness(v_4)/1 = 100$
    (d) Do you expect the number of strings matching S to increase, decrease, or stay the same in subsequent generations? Explain your answer! [4] *As S matches exactly one of the current strings and that string has maximum fitness and so is guaranteed to survive due to the elitism except in the unlikely event where a whole population of maximum fitness individuals is evolved, the number will not decrease; however, considering its maximum order and maximum defining length, the fact that genotypically similar individuals have much lower fitness, and the absence of any other high fitness individuals in the current population, it is questionable whether any clones will appear, so increasing seems somewhat tenuous and most likely it will stay the same till termination.*

20. Say you want to purchase a new house and care most about maximizing space and affordability. You collect square footage data and pricing on ten different houses and then you normalize both the square footage data and the pricing which results in the following table, where higher space numbers indicate greater square footage and higher affordability numbers indicate lower prices:

ID	Space	Affordability
1	5	7
2	2	6
3	3	2
4	4	10
5	7	8
6	5	6
7	3	4
8	10	2
9	1	1
10	6	1

(a) List for each element which elements it dominates; indicate elements with their IDs. [4 pts]

ID	Dominates
1	2,3,6,7,9
2	9
3	9
4	2,3,7,9
5	1,2,3,6,7,9,10
6	2,3,7,9
7	3,9
8	3,9,10
9	None
10	9

(b) Show the population distributed over non-dominated levels, like some multi-objective EAs employ, after each addition of an element, starting with element 1 and ending with element 10, increasing the element number one at a time; indicate elements with their IDs. So you need to show ten different population distributions, the first one consisting of a single element, and the last one consisting of ten elements. [12 pts]

After adding element 1:

Level 1: 1

After adding element 2:

Level 1: 1

Level 2: 2

After adding element 3:

Level 1: 1

Level 2: 2,3

After adding element 4:

Level 1: 1,4

Level 2: 2,3

After adding element 5:

Level 1: 4,5

Level 2: 1

Level 3: 2,3

After adding element 6:

Level 1: 4,5

Level 2: 1

Level 3: 6

Level 4: 2,3

After adding element 7:

Level 1: 4,5

Level 2: 1

Level 3: 6

Level 4: 2,7

Level 5: 3

After adding element 8:

Level 1: 4,5,8

Level 2: 1

Level 3: 6

Level 4: 2,7

Level 5: 3

After adding element 9:

Level 1: 4,5,8

Level 2: 1

Level 3: 6
Level 4: 2,7
Level 5: 3
Level 6: 9

After adding element 10:

Level 1: 4,5,8
Level 2: 1,10
Level 3: 6
Level 4: 2,7
Level 5: 3
Level 6: 9

21. Given the following two parents with permutation representation:
$p1 = (475318692)$
$p2 = (524836971)$

(a) compute the first offspring with Cycle Crossover. Show first the cycles you've identified and then the construction of the offspring. [6 pts]

Cycle 1: 4-5, Cycle 2: 7-2-1-3-8-6-9

Construction of first offspring by scanning parents from left to right, starting at parent 1 and alternating parents:

i. Add cycle 1 from parent 1: $4 \cdot 5 \cdots\cdots$
ii. Add cycle 2 from parent 2: 425836971

(b) Compute the first offspring with PMX, using crossover points between the 2nd and 3rd loci and between the 6th and 7th loci. Show your offspring construction steps. [6 pts]

i. $\cdot \cdot 5318 \cdot \cdot\cdot$
ii. $4 \cdot 5318 \cdot \cdot\cdot$
iii. $4 \cdot 5318 \cdot \cdot 6$
iv. 425318976

(c) Compute the first offspring with Order Crossover, using crossover points between the 3rd and 4th loci and between the 7th and 8th loci. Show your offspring construction steps. [4 pts]

i. Child 1: $\cdot\cdot\cdot 3186 \cdot \cdot$
ii. Child 1: **249318675**

(d) Compute the first offspring with Edge Crossover, except that for each random choice you instead select the lowest element. Show how you arrived at your answer by filling the following templates: [12 pts]

Edge Table:

Element	Edges

Construction Table:

Element Selected	Reason Selected	Partial Result

Original Edge Table:

Element	Edges	Element	Edges
1	3,8,7,5	6	8,9+,3
2	9,4+,5	7	4,5,9,1
3	5,1,8,6	8	1,6,4,3
4	2+7,8	9	6+,2,7
5	7,3,1,2		

Construction Table:

Element selected	Reason	Partial result
1	Lowest	1
3	Equal list size, so lowest	13
5	Equal list size, so lowest	135
2	Equal list size, so lowest	1352
4	Common edge	13524
7	Equal list size, so lowest	135247
9	Only element	1352479
6	Only element	13524796
8	Last element	**135247968**

Edge Table After Step 1:

Element	Edges	Element	Edges
1	3,8,7,5	6	8,9+,3
2	9,4+,5	7	4,5,9
3	5,8,6	8	6,4,3
4	2+7,8	9	6+,2,7
5	7,3,2		

Edge Table After Step 2:

Element	Edges	Element	Edges
		6	8,9+
2	9,4+,5	7	4,5,9
3	5,8,6	8	6,4
4	2+7,8	9	6+,2,7
5	7,2		

Edge Table After Step 3:

Element	Edges	Element	Edges
		6	8,9+
2	9,4+	7	4,9
		8	6,4
4	2+7,8	9	6+,2,7
5	7,2		

Edge Table After Step 4:

Element	Edges	Element	Edges
		6	8,9+
2	9,4+	7	4,9
		8	6,4
4	7,8	9	6+,7

Edge Table After Step 5:

Element	Edges	Element	Edges
		6	8,9+
		7	9
		8	6
4	7,8	9	6+,7

Edge Table After Step 6:

Element	Edges	Element	Edges
		6	8,9+
		7	9
		8	6
		9	6+

Edge Table After Step 7:

Element	Edges	Element	Edges
		6	8
		8	6
		9	6+

22. Alice is writing an EA to solve the binary knapsack constraint satisfaction problems where the sum of the item costs exceeds the constraint limit. Recall that this problem consists of identifying a set of items that maximize value while keeping their cumulative cost below a known limit. Given the following constraint handling approaches:

    (a) Ignore the constraints under the motto: all is well that ends well.

    (b) Upon generating an infeasible solution, immediately kill it and generate a new solution; repeat this step until a feasible solution is generated.

    (c) Employ a penalty function that reduces the fitness of infeasible solutions, preferably so that the fitness is reduced in proportion to the number of constraints violated, or to the distance from the feasible region.

    (d) Employ a repair function that takes infeasible solutions and "repairs" them by transforming them into a related feasible solution, typically as close as possible to the infeasible one.

    (e) Employ a closed feasible solution space which guarantees that the initial population consists of feasible solutions only and all evolutionary operations on feasible solutions are guaranteed to result in feasible solutions. Typically a combination of custom representation, initialization, recombination, and mutation is employed to achieve this.

    (f) Employ a decoder function that maps genotype space to phenotype space such that the phenotypes are guaranteed to be feasible even when the genotypes are infeasible. Typically this involves mapping multiple different genotypes to the same phenotype.

    Which (combination) of these six constraint handling approaches do you recommend Alice employs? Explain your answer! [5 pts]

    If Alice knows that the ratio of invalid to total solutions is very low, for instance if the sum of the item costs doesn't exceed the constraint limit by much, then use approach (b) where invalid solutions are immediately discarded and use either stochastic survival or a mutation with for instance a Gaussian distributed mutation rate to guarantee global optimum reachability. Otherwise use approach (f) with a high quality decoder function which guarantees valid solutions while imposing no limitations on the search of the genotype space.