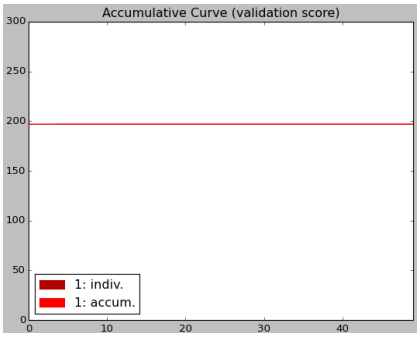
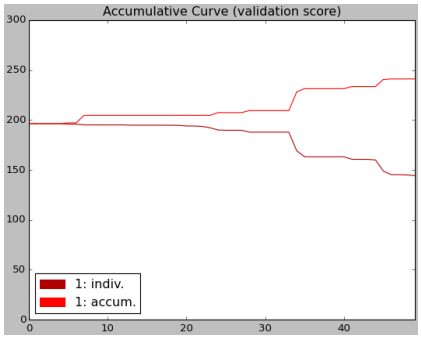
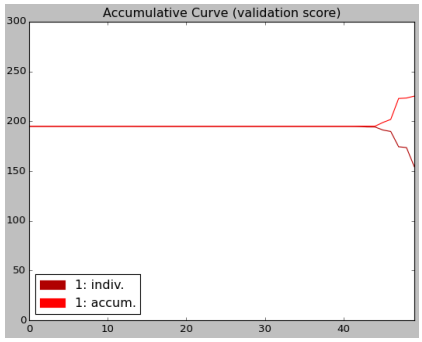
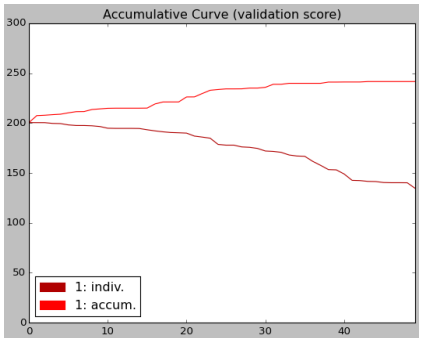
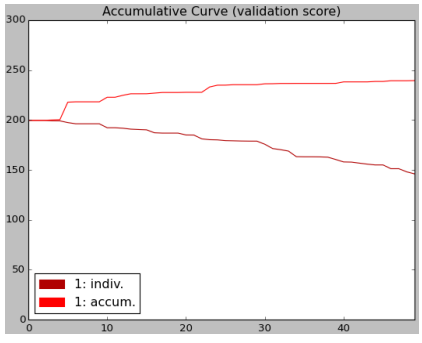
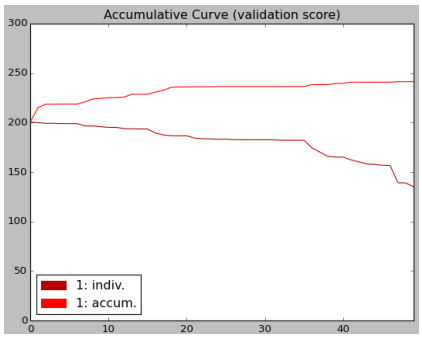
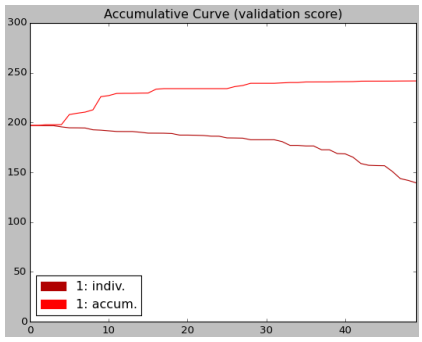
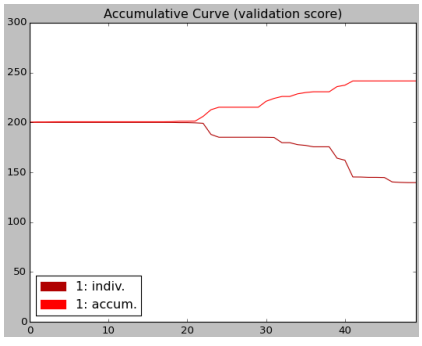
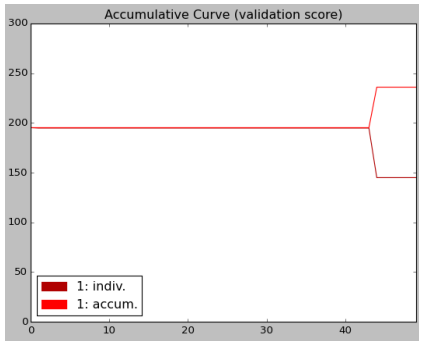
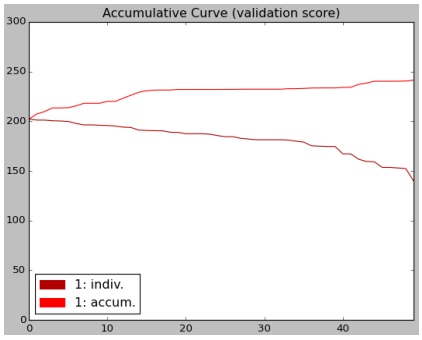
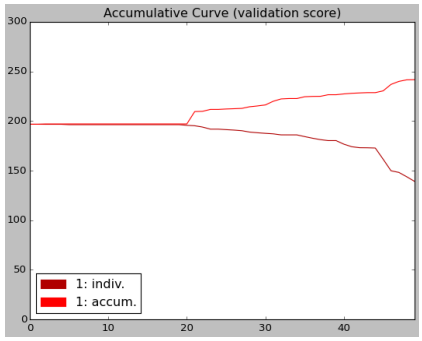
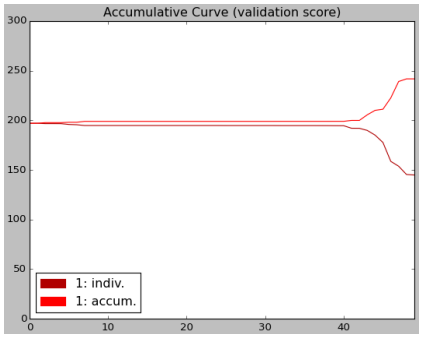
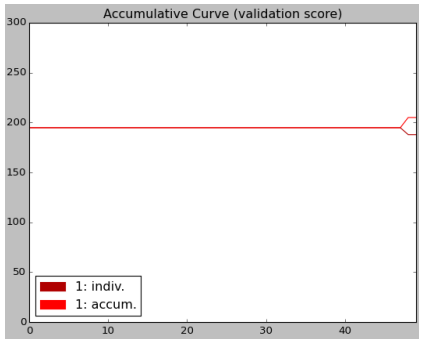
Accumulative Curve for Various Diversity Metrics and Codings

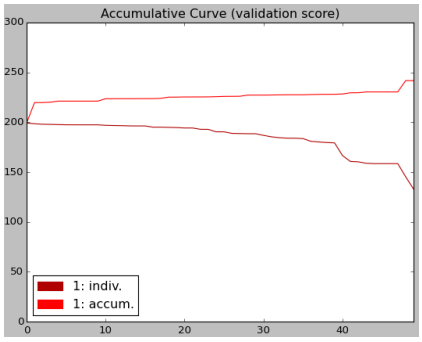
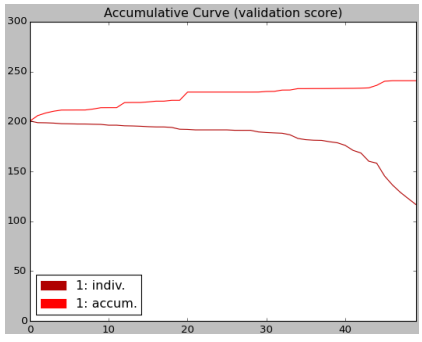
Example of coding schema for one hand played:
C1: seed+position+[board_cards+action(as letter)]+, eg.: 241206h2sQsc6h2sQsAhr
C2: [action(as number)]+, eg.: 1111122220
C3: quantized aggressiveness, eg.: 2
C4: [quantized HS+quantized EP+action(as letter)]+, eg.: 24c34c44r
Obs.: C3 and C4 were tested using quantization for 3 groups and for 5 groups.

NO CODE	No Diversity	Genotype	Fitness Sharing
			

C1 / C2	NCD (C1)	Entropy (C2)
		

C3, 3 groups	NCD (C3, 3 groups)	Hamming (C3, 3 groups)
		
	Euclidean (C3, 3 groups)	**Entropy (C3, 3 groups)**
		

C3, 5 groups	NCD (C3, 5 groups)	Hamming (C3, 5 groups)
		
	Euclidean (C3, 5 groups)	**Entropy (C3, 5 groups)**
		

C4	NCD (C4, 3 groups)	NCD (C4, 5 groups)
		

Conclusions:

- best ones: **NCD (C1), Entropy (C2), NCD (C3, 3 groups), Hamming (C3, 3 groups), NCD (C3, 5 groups), NCD (C4, 3 groups), NCD (C4, 5 groups)**
- Question1: Are they good enough, or still need to be improved?
  - If they are good enough, then I will compare them using other metrics to select the overall best one or ones, and then start the implementation of the second layer of SBB.
  - Else, I will try to improve the best ones to find better diversity metrics.
- Question2: What are other metrics that I could use to ensure they are indeed promoting diversity?
  - I also looked if the fitness, validation, and champion score were improving over time (instead of getting stuck), and if the diversity curve was steady. From these metrics I removed the **Entropy (C2)** from the finalists (nearly random global diversity and very unstable global fitness).
- Question3: Are this set of results relevant enough to add to my thesis or to a paper?

Next step:

- Since the other six diversities didn't get stuck performance scores, I am performing longer runs with them to find out at what point they get stuck/if they get stuck.
- While I am executing the runs I am organizing and structuring the content for the literature review chapter.