SP Assignment 3: Points, Points, and more Points

USN: 303039534

1 Introduction

2 Results

We start by implementing the Dmin model as described in the assignment and plot an example of a dmin simulation with parameters n=200, m=30, s=5 and ranges of x=[200:100] and y = [100:900]
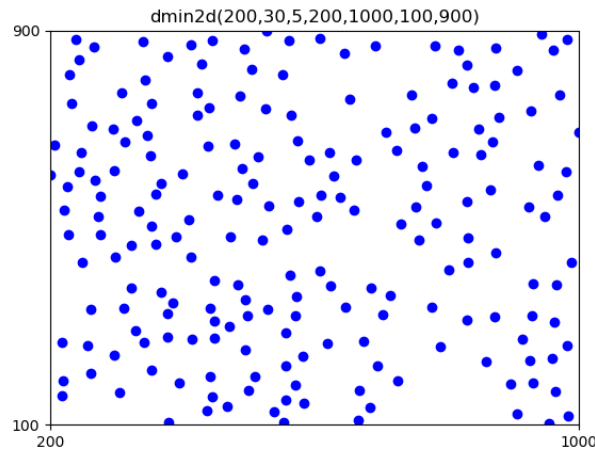

Figure 1: dmin model

We proceed to implement a function calc_RI to calculate the regularity index of a given pattern. The regularity index is defined as

$$RI = \frac{\text{mean}(d_{ij}}{stdev(d_{ij})} \tag{1}$$

where d_{ij} is the distance between points i and j. This tells us how regular the spacing of points is; if all points have pairwise similar spacings, the standard deviation will be low and the regularity index high. If, on the other hand, pairwise distances differ a lot, the standard deviation will be high and the regularity index low. This is weighted by the mean distance to account for scalings of the grid.

We then generate 1000 dmin patterns with n=200, m=0, s=0, x=[0:1000], y=[0:1000]; i.e. we generate a square pattern of 200 points with no exclusion zone. We report the 50th largest value as a measure of regularity for this set of parameters. This is the boundary between the 5% highest and 95% lowest regularity indices we can achieve given this set of parameters, and thus tells us how regular a pattern we can expect to generate within a reasonable number of trials (here 20).

This gives a result of xxx. We can repeat this calculation 10 times which gives us a mean of xxx and an std of yyy indicating something.

We can investigate how this regularity index changes with both the number of points added and the shape of the grid. We constrain the area of the grid to $10\hat{6}$ giving us one degree of freedom specifying the shape, and we quantify this by the ratio of the two dimensions. We scan this ratio as well as the number of points added and report.

reproduce normal data. Start by scanning:

see that it's roughly 12 and 3; set niter=1000 - get sd = 0.11

Thus write steepest descent with threshold 0.2 (due to uncertainty in slope)

map line onto 3d plot

square packing: 21*21 = 441

separation: sqrt(20$\hat{2}$-10$\hat{2}$) hexagonal: floor(400/sqrt(300)) = 23; i.e. 24 rows 12*(21+20) = 12*41 = 492
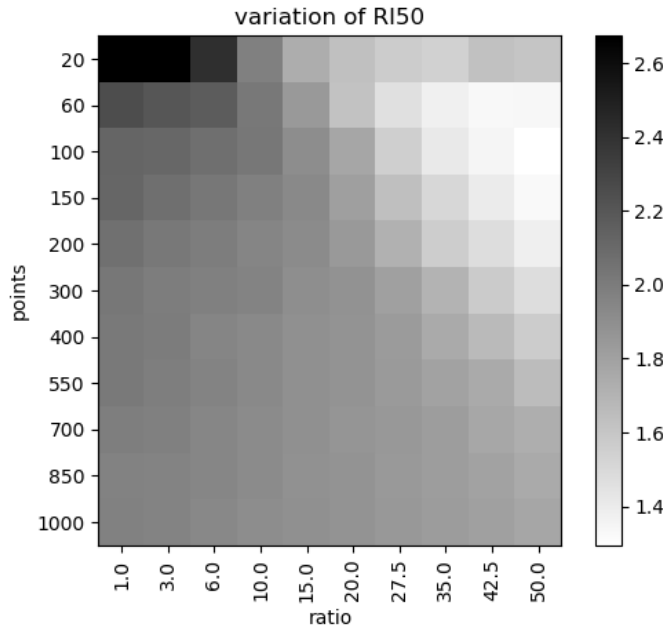
Figure 2: dependence of RI50 on n and ratio

H.sapiens	S249	T252	T356	T373	S608	S612	S780	S788	S795	S807	S811	T821	T826
C.jacchus	S	T	T	T	S	S	S	S	S	S	S	T	T
M.mulatta	S	T	T	T	S	S	S	S	S	S	S	T	T
F.catus	S	T	T	T	S	S	S	S	S	S	**N**	T	T
M.musculus	S	T	T	T	S	S	S	S	S	S	S	T	T
R.norvegicus	S	T	T	T	S	S	S	S	S	S	S	T	T
G.gallus	S	T	T	T	S	S	S	S	S	S	S	**S**	T
M.gallopavo	S	T	T	T	S	S	S	S	S	S	S	**S**	T
I.punctatus	S	T	T	T	S	-	S	S	S	S	S	-	T
D.rerio	S	T	T	T	S	-	S	S	S	S	S	-	T
RBL1	F	L	T	Q	M	P	P	I	-	I	P	G	-
RBL2	-	N	K	P	R	P	P	V	A	I	Q	A	F

Table 1: Conservation of phosphorylation sites. Mismatches in orthologs have been highlighted in bold

3 Discussion & Conclusion

4 Methods

5 References

Chinnam & Goodrich: "RB1, development, and cancer". *Curr Top Dev Biol.* 94 (2013)
Dyson: "*RB1*: "a prototype tumor supressor and an enigma", *Genes and Development* 30 (2016)
Rubin: "Deciphering the Rb1 phosphorylation code", *Trends Biochem Sci.* 38 (2013)
Narasimha et al.: "Cyclin D activates the Rb tumor suppressor by mono-phosphorylation", *eLife* 3 (2014)
Turro et al.: "Haplotype and isoform specific expression estimation using multi-mapping RNA-seq reads", *Genome Biology* 12 (2013)
Pearson & Lipman: "Improved tools for biological sequence comparison", *Proc Natl Acad Sci USA* 85 (1998)
Yam et al.: "Cyclin A in cell cycle control and cancer", *Cell. Mol. Life Sci.* 59 (2002)
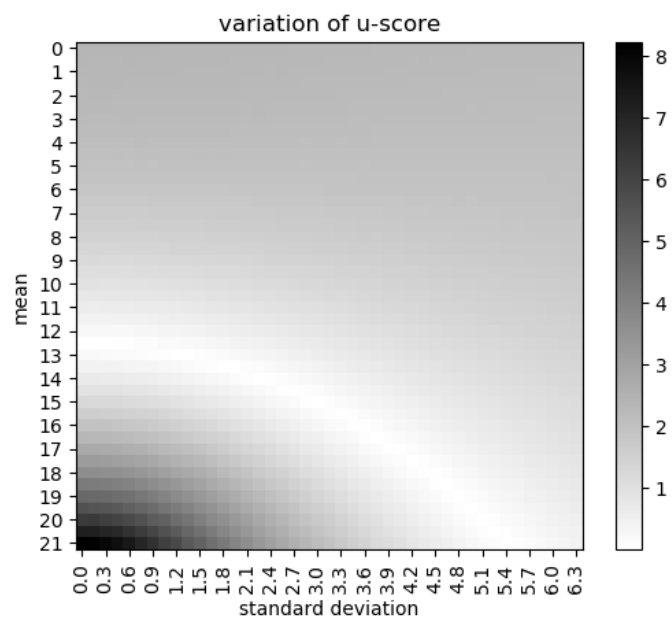
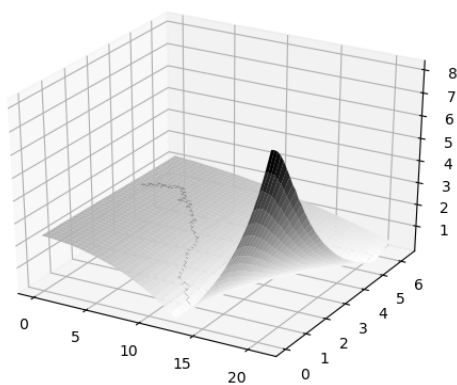## 6 Appendix (code)

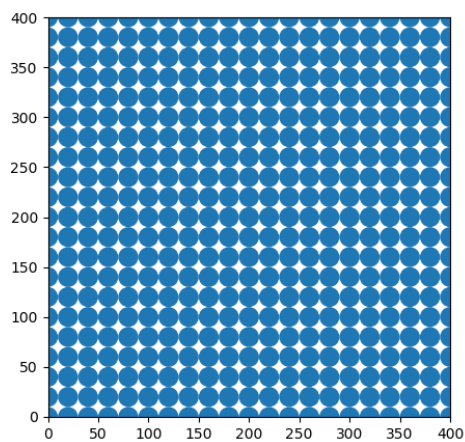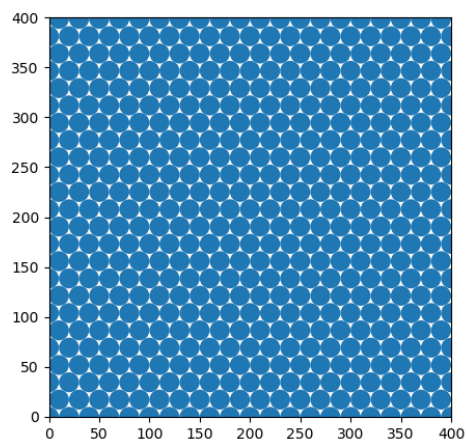Figure 3: similarity for different parameters
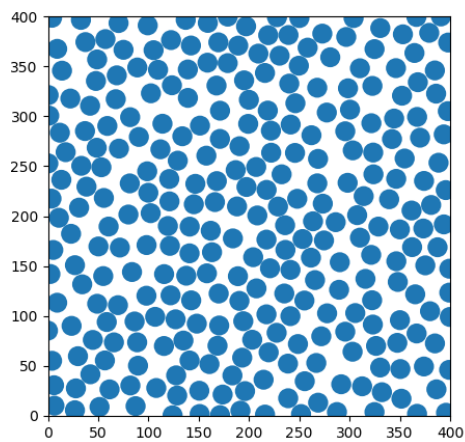

Figure 4: similarity for different parameters
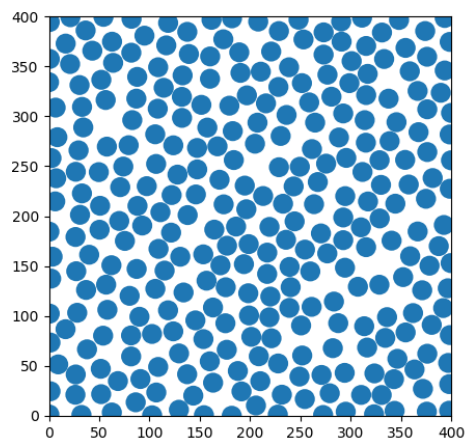
(a)

(b)

Figure 5: different packings.

(a) 294 points generated by the sequential algorithm

(b) 302 points generated by the epoch algorithm

Figure 6: different packings.

4