

Algorithms in Genome Research  
Winter 2012/2013

Exercises

Number 4, Discussion: 16. November 2012

1. Give an example of perfectly balanced binary search tree storing 8 (value,key) pairs in its leaves as described in Lemma 2.2.6. Give an example of a range minimum query for some non-empty interval.
2. A *van Emde Boas tree* (vEB tree) supports in  $O(\log \log n)$  time insertions, deletions, and *predecessor queries* for values in interval  $[1, n]$ . Predecessor query returns the largest element  $i'$  stored in the vEB tree smaller than query element  $i$ . Show how the structure can be used instead of a balanced search tree of Lemma 2.2.6 to solve range minimum queries for semi-infinite intervals  $(-\infty, i]$  (i.e. for the type of queries we used e.g. in the LCS algorithm).
3. SOLiD<sup>1</sup> sequencing produces short reads of DNA in *colour-space* with a two-base defined by the matrix (row=first base, column=second base):

	A	C	G	T
A	0	1	2	3
C	1	0	3	2
G	2	3	0	1
T	3	2	1	0

For example, T012023211202102 equals TTGAAGCTGTCCTGGA (first base is always given). Modify overlap alignment to work properly in the case where one of the sequences in SOLiD read and the other is a normal sequence.

---

<sup>1</sup>TM Applied Biosystem