Libpll sequential benchmarks

November 30, 2016

1 Benchmark description

The following benchmarks compare several libpll implementations with different modes. They measure the execution time of a full likelihood computation on a fixed tree. To avoid measuring the initialization part, we repeat several times `pll_update_partials` and `pll_compute_edge_likelihood` on the same partitions and tree.

- xflouris means that the implentation used is this one: https://github.com/xflouris/libpll.
- bmorel means that the implementation used is this one: https://github.com/BenoitMorel/libpll. It supports sites repeats, and the data structure used is a bit different from xflouris (even without sites repeats): CLVs are not assumed to be sorted by sites, and an additional lookup table is used to access them in most of the core functions.
- bmorel2 is a (temporary) hacked version of bmorel where I do not use the new lookup structure for the scalers, because I think it slows down the execution, and the speed up with the sites repeats is not great.
- default mode means that the option `PLL_ATTRIB_PATTERN_TIP` and `PLL_ATTRIB_SITES_REPEATS` are unset.
- tip pattern means that the option `PLL_ATTRIB_PATTERN_TIP` is set.
- sites repeats means that the option `PLL_ATTRIB_SITES_REPEATS` is set.
- M is the size of the buffer allocated to compute the sites repeats class identifiers. When it increases, more nodes can benefit from sites repeats.

2 Summary of the results

Without sites repeats, bmorel's implementation with the additional lookup table can be from 0% to 10% slower than xflouris' one, especially with SSE and AVX architectures. bmorel2's hacked implementation (so without using the lookup for the scalers) is better, but can also be slower than xflouris'one.

In all versions, the tip pattern mode is around 1.5 faster than the default mode.

The sites repeats speed up is between 2 and 5 times faster than the tip pattern mode. It depends on :

- the dataset
- the architecture : it performs a bit better with CPU than with AVX and SSE
- the size of the sites repeats matrix buffer (M) : it performs better when the size increases

3 Benchmark

CPU architecture, 500 iterations

	seq59	seq128	seq404
xflouris tip pattern	5803.58 ms	77073.3 ms	96565.4 ms
bmorel tip pattern	6015.05 ms	78180.2 ms	99699 ms
bmorel sites repeats M=1000	4512.74 ms	48634.2 ms	69984.5 ms
bmorel sites repeats M=10000	3590.54 ms	37765.8 ms	48610 ms
bmorel sites repeats M=1000000	2638.64 ms	33579.4 ms	23267.2 ms
bmorel sites repeats no update sr M=1000000	2010.07 ms	21392 ms	16036.4 ms

SSE architecture, 500 iterations

	seq59	seq128	seq404
xflouris tip pattern	2493.25 ms	34267.9 ms	39762.5 ms
bmorel tip pattern	2743.76 ms	36185.4 ms	43582.2 ms
bmorel sites repeats M=1000	1881.72 ms	23373.3 ms	30213.8 ms
bmorel sites repeats M=10000	1658.64 ms	18513.3 ms	23476.5 ms
bmorel sites repeats M=1000000	1413.62 ms	15355.4 ms	15024.1 ms
bmorel sites repeats no update sr M=1000000	880.667 ms	9850.06 ms	7115.65 ms

AVX architecture, 500 iterations

	seq59	seq128	seq404
xflouris tip pattern	2277.22 ms	33698.9 ms	35714.6 ms
bmorel tip pattern	2411.97 ms	35418.5 ms	37865.3 ms
bmorel sites repeats M=1000	1730.36 ms	20658.5 ms	26186.1 ms
bmorel sites repeats M=10000	1475.74 ms	17745 ms	20227.1 ms
bmorel sites repeats M=1000000	1326.07 ms	14933.5 ms	13366.4 ms
bmorel sites repeats no update sr M=1000000	831.47 ms	8907.74 ms	6520.19 ms