Appendix to "A Common Framework for Linear and Cyclic Multiple Sequence Alignment Problem"

Sebastian Will   Peter F. Stadler

1 Test instances and optimal alignments

Instance 1

3 sequences of length 10

```
>a
CGGCAUCGGC
(([[..))]]
>b
CGACCUCGGG
((.[[.))]]
>c
GCGCUGACGC
(([[.).)]]
```

Optimal alignment

```
a       CG-GCAUC-GGC
a       ((-[[..)-)]]
b       CGACC-UC-GGG
b       ((.[[-.)-)]]
c       GC-GC-UGACGC
c       ((-[[-.).)]]
```

Instance 1R

3 circular sequences of length 10; sequences of instance 1 circularized and cut at non-homologous positions

```
>a
CGGCAUCGGC
(([[..))]]
>b
```

```
ACCUCGGGCG
.[[.((]]))
>c
UGACGCGCGC
.(.([[))]]
```

Optimal circular alignment, which was found in our experiment within the time limit. (By design this is equivalent to the optimal solution of instance 1.)

```
a       CG-GCAUC-GGC
a       ((-[[..)-)]]
b       CGACC-UC-GGG
b       ((.[[-.)-)]]
c       GC-GC-UGACGC
c       ((-[[-.).)]]
```

1.1   Instance 2

3 sequences of length 15

```
>a
CGGCAAAGCCGGCGC
((([[...{{))]]}}
>b
CGGCACGACGUGCCG
((([.{{.)).]]}}
>c
GGUCGUGCACCCGGC
((.[[.{{.))]]}}
```

Optimal alignment

```
a       CG-GCAAAGC-CG-GCGC
a       ((-[[...{{-))-]]}}
b       CG-GC-A-CGACGUGCCG
b       ((-[[-.-{{.)).]]}}
c       GGUCG-U-GCACC-CGGC
c       ((.[[-.-{{.))-]]}}
```

1.2   Instance 3

3 sequences of length 20

```
>a
CGCGCGAAAGCCGCGCGAGC
((((([[...{{)))]]].}}
>b
CGCGCGACGACGCUGCGCGA
```

```
(((([[[.{{.))).]]]}}.
>c
AGGGUCGCUAGCACCCGCGGC
.(((.[[[..{{.)))]]]}}
```

Optimal alignment

```
a       -CGC-GCGAAAGC-CGC-GCGAGC-
a       -(((-[[[...{{-)))-]]].}}-
b       -CGC-GCG--ACGACGCUGCG-CGA
b       -(((-[[[--.{{.))).]]]-}}.
c       AGGGUCGCU-AGCACCC-GCG-GC-
c       .(((.[[[.-.{{.)))-]]]-}}-
```

1.3 Instance 4

4 Sequences of length 10

```
>a
CGGCAUCGGC
((([[..))]]
>b
CGACCUCGGG
((.[[.))]]
>c
GCGCUGACGC
((([[.).)]]
>d
CGGGACGAGG
((([[.)).]]
```

Optimal alignment

```
a       CG-GCAUC-G-GC
a       ((-[[..)-)-]]
b       CGACC-UC-G-GG
b       ((.[[-.)-)-]]
c       GC-GC-UGAC-GC
c       ((-[[-.).)-]]
d       CG-GGA-C-GAGG
d       ((-[[.-)-).]]
```

1.4 Instance 4R

4 circular sequences of length 10; sequences of instance 4 circularized and cut at non-homologous positions

```
>a
CGGCAUCGGC
(((([..))]]
>b
ACCUCGGGCG
.[[.((]]))
>c
UGACGCGCGC
.(.([[))]]
>d
GAGGCGGGAC
(.[[)(]].)
```

Optimal alignment (identical to optimal alignment of Instance 4)

```
a       CG-GCAUC-G-GC
a       ((-[[..)-)-]]
b       CGACC-UC-G-GG
b       ((.[[-.)-)-]]
c       GC-GC-UGAC-GC
c       ((-[[-.).)-]]
d       CG-GGA-C-GAGG
d       ((-[[.-)-).]]
```