

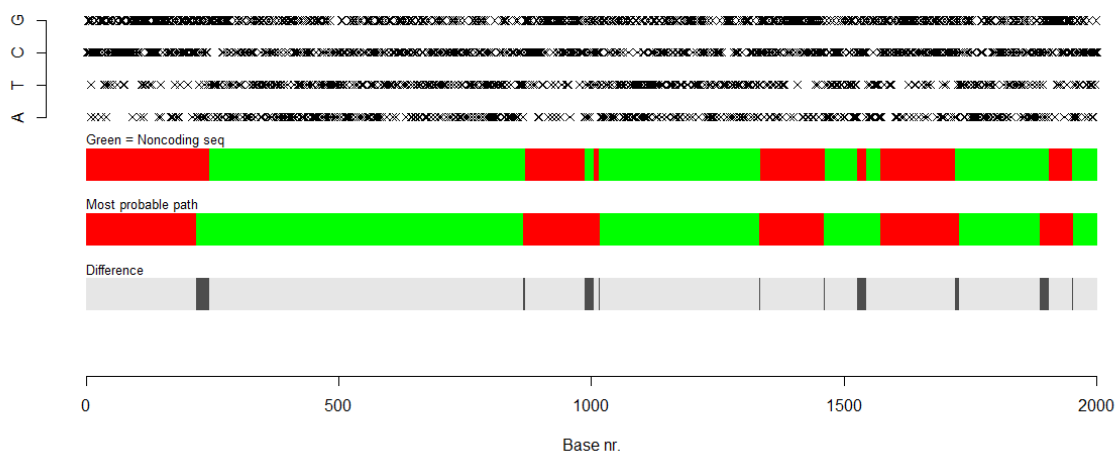
2) Das Programm erzeugt mit festgelegten Wahrscheinlichkeiten Würfelergebnisse, die zwischen einem gezinkten und einem fairen Würfel changieren. Anschließend wertet das Programm die Ergebnisse blind mit einem Hidden Markov Model dahingehend aus, welche Sequenzen mit dem fairen und welche mit dem gezinkten Würfel erzeugt wurden. Beide Ergebnisse, sowie die Differenz zwischen beiden wird grafisch dargestellt.

3)

```
transProbs = matrix(c(0.99, 0.01, 0.01, 0.99)
```

```
emissionProbs = matrix(c(rep(1/4, 4), c(0.1, 0.1, 0.4, 0.4)
```

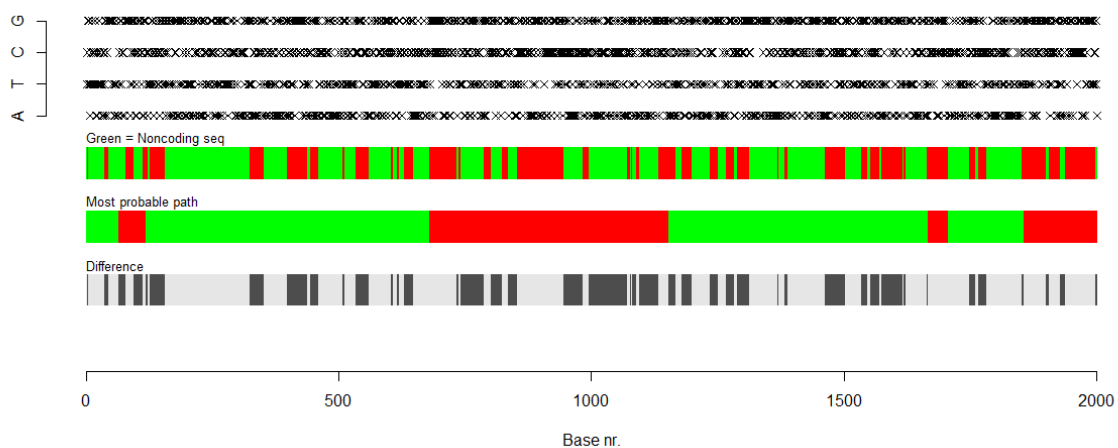
Noncoding and coding seq



```
transProbs = matrix(c(0.95, 0.05, 0.05, 0.95)
```

```
emissionProbs = matrix(c(rep(1/4, 4), c(0.15, 0.15, 0.35, 0.35)
```

Noncoding and coding seq



Im ersten Fall wurde mit einer geringen Transitions Wahrscheinlichkeit gearbeitet und einer hohen Emissions Wahrscheinlichkeit für C/G-Basen. Das HMM liefert ein vergleichsweise gutes Modell für die zugrundeliegenden Wahrscheinlichkeiten. Fehler entstehen häufig bei sehr kurzen Sequenzen, die in zwei längere gegenteilige Sequenzen eingebettet sind.

Im zweiten Fall wurde die Transitions Wahrscheinlichkeit erhöht und die Emissions Wahrscheinlichkeit verringert, das Programm arbeitet nun deutlich unzuverlässiger. Gerade die Erhöhung der Transitions Wahrscheinlichkeit erhöht die Differenz zwischen tatsächlichem Algorithmus und HMM stark.