Adaptive Huffman Coding

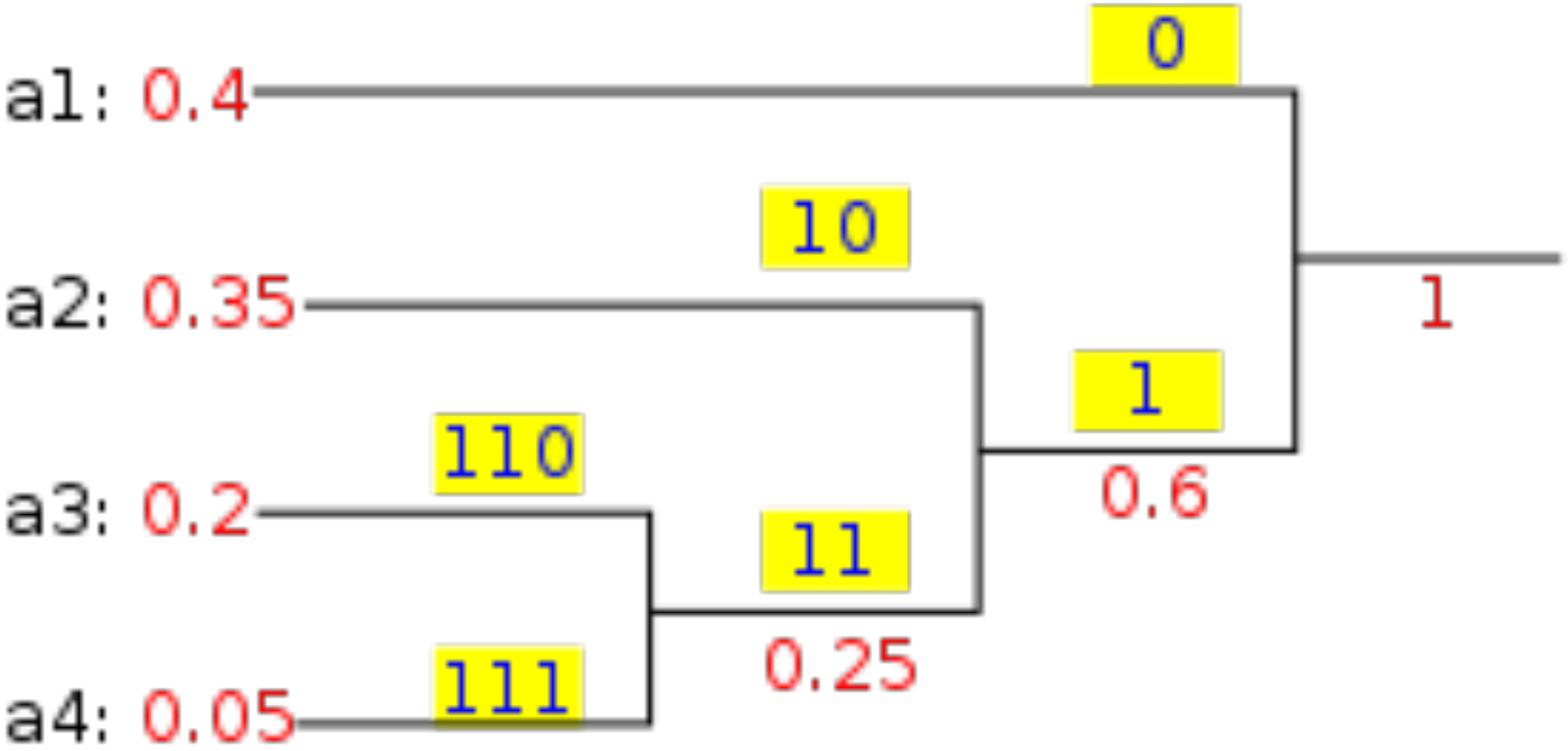
Martin Matyášek, Tomáš Hnídek

Huffman Coding

- lossless data compression, prefix code
- runs in O(n*log(n))
- tree construction algorithm:
  1. foreach symbol s do add(queue, node(s))
  2. while not empty(queue) do
     i. n1 := poll(queue) n2 := poll(queue)
     ii. n := new node with children n1, n2 and p(n) = p(n1) + p(n2)
     iii. add(queue, n)
  3. remaining node is the root (the tree has been generated)

Huffman Coding

Example of the coding tree:



Adaptive Huffman Coding

- lossless data compression, prefix code
- adaptive!
- runs in O(n*log|Λ|), space O(|Λ|) where
  - n - size of the input message
  - Λ - the alphabet
- algorithms:
  - FGK (Faller 1973, Gallager 1978, Knuth 1985)
  - Vitter's algorithm (1987)

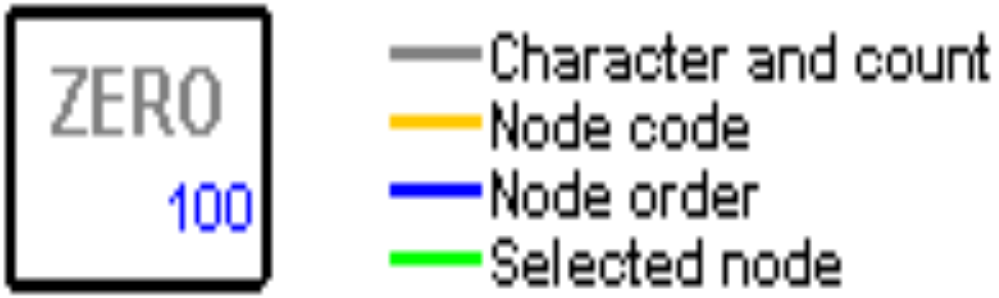
Adaptive Huffman Coding

- definition:
  - binary coding tree has the sibling property if each node (except root) has a sibling and if the nodes can be listed in order of nonincreasing weight with each node adjacent to its sibling
- theorem:
  - binary prefix code is Huffman code iff the code tree has the sibling property
- main idea:
  - modify code tree so that it satisfies the sibling property

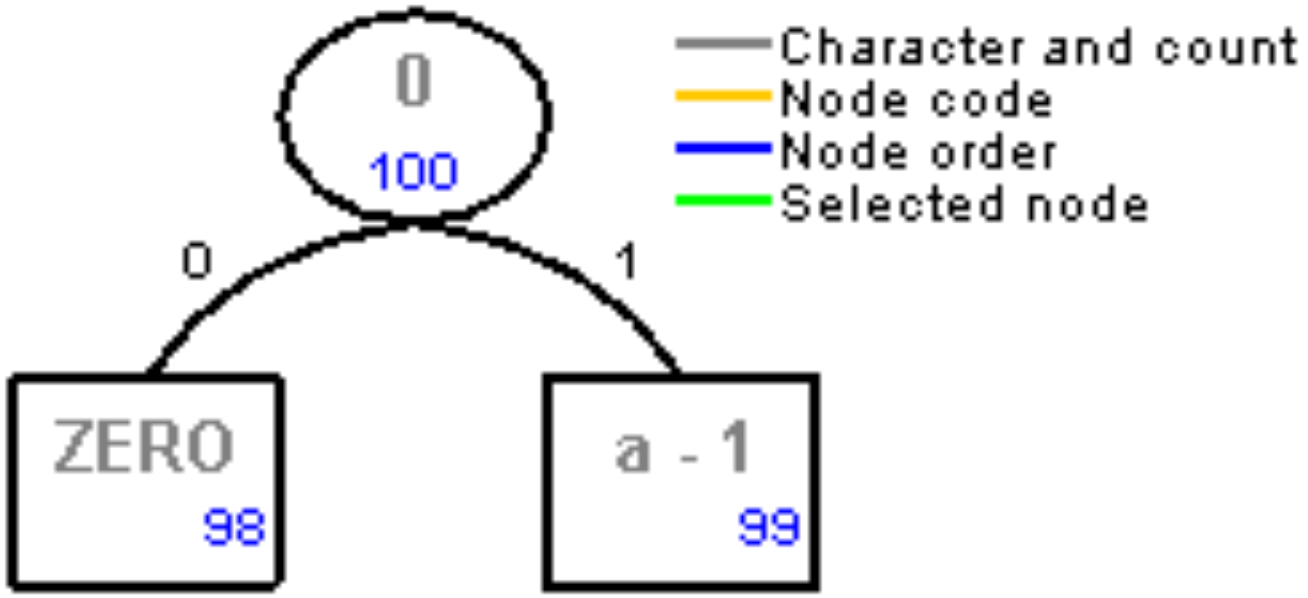
FGK algorithm

- pros:
  - can compress better than static Huffman coding, due to the fact that coding tree needn't to be encoded
- cons:
  - has worse sensitivity to errors, because one wrong character can destroy whole message

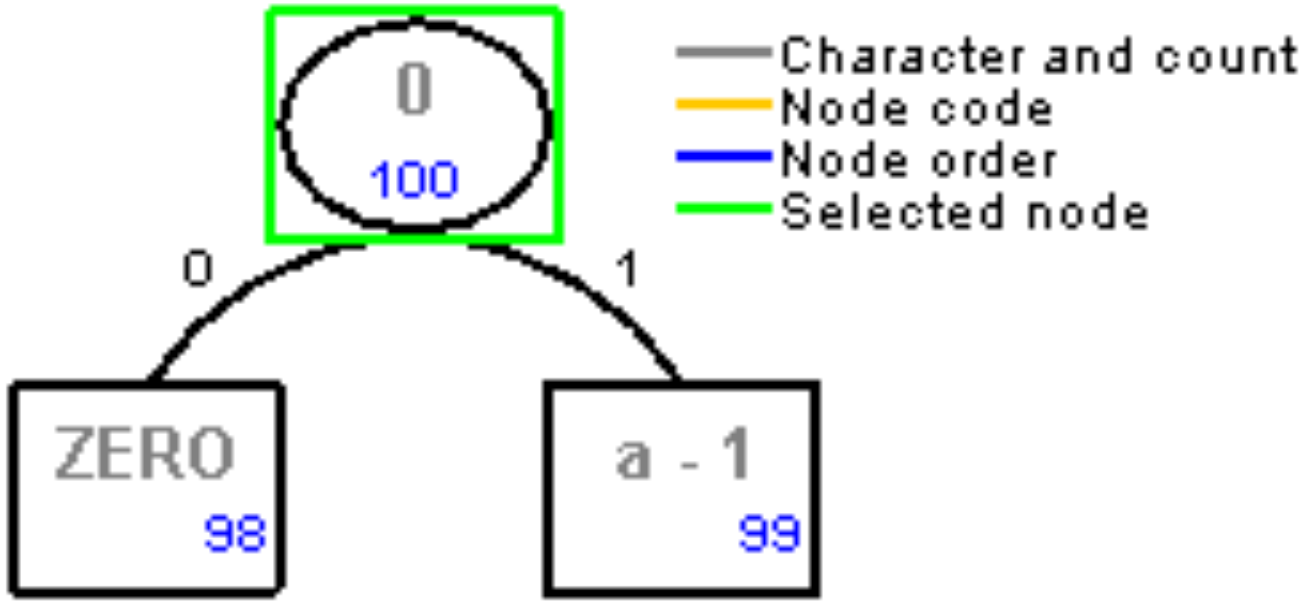
FGK Example - abba



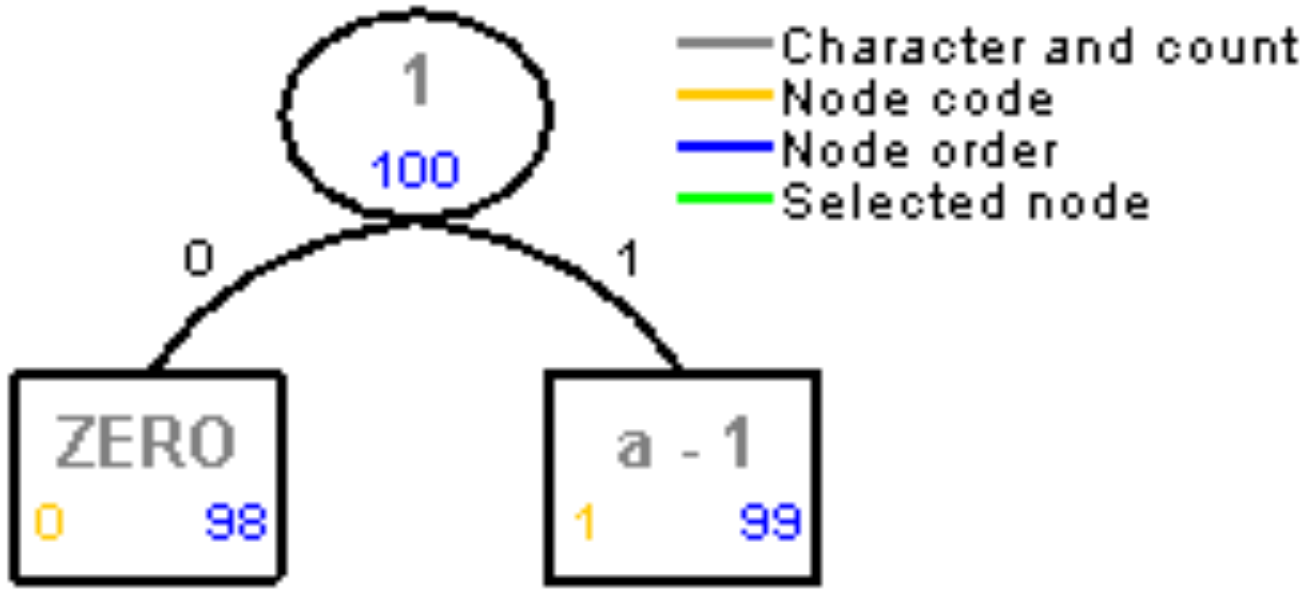
FGK Example - abba



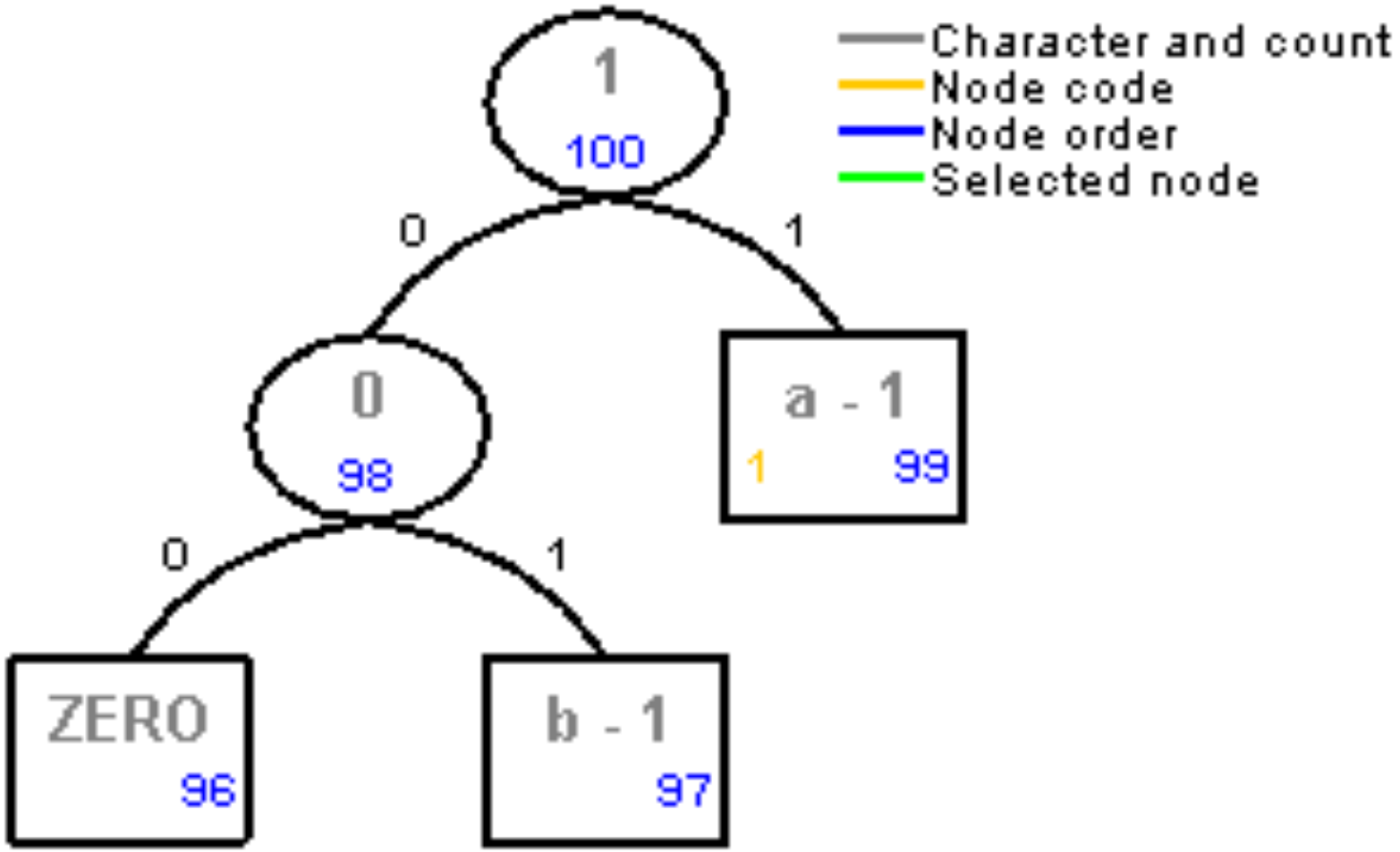
FGK Example - abba



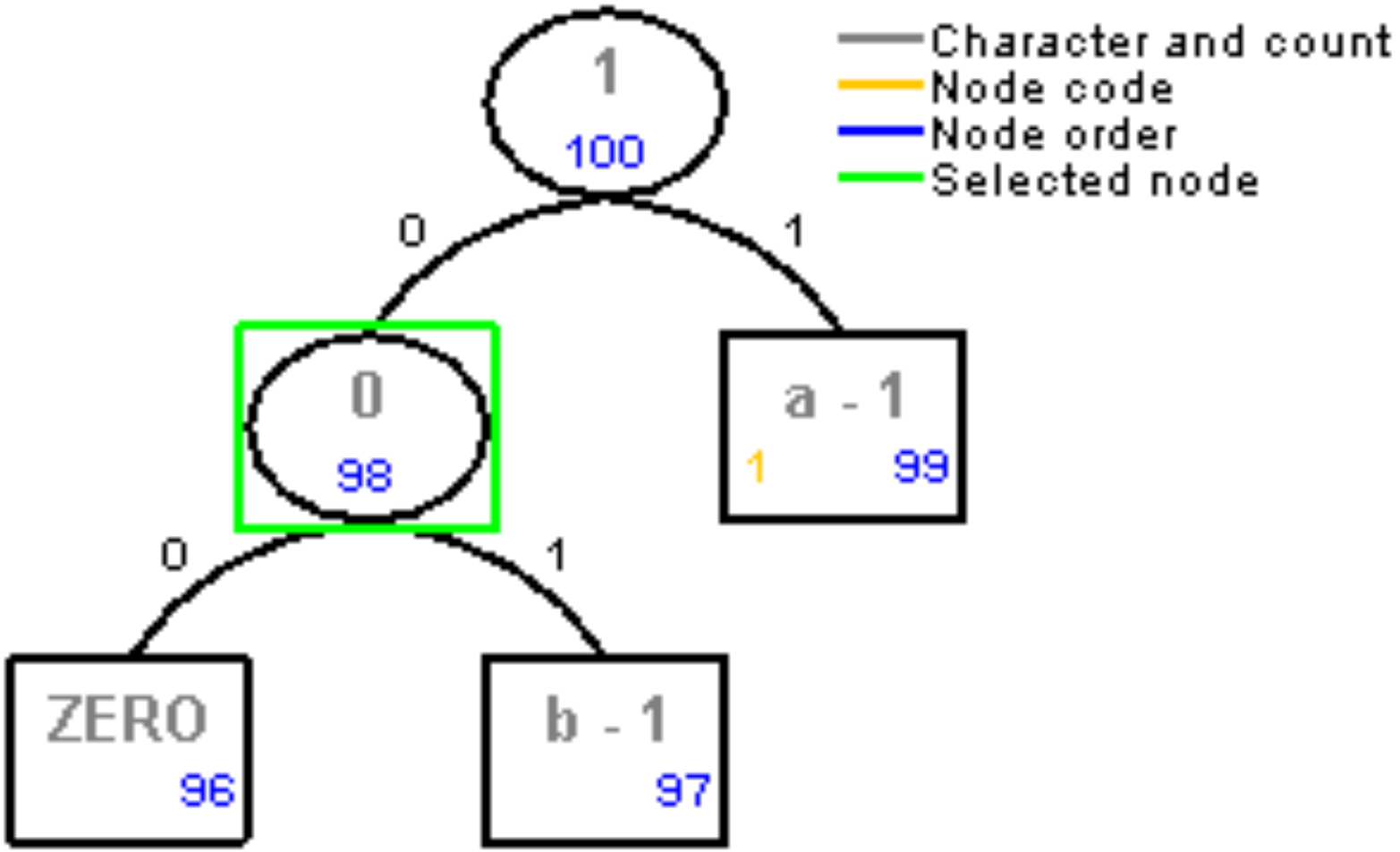
FGK Example - abba



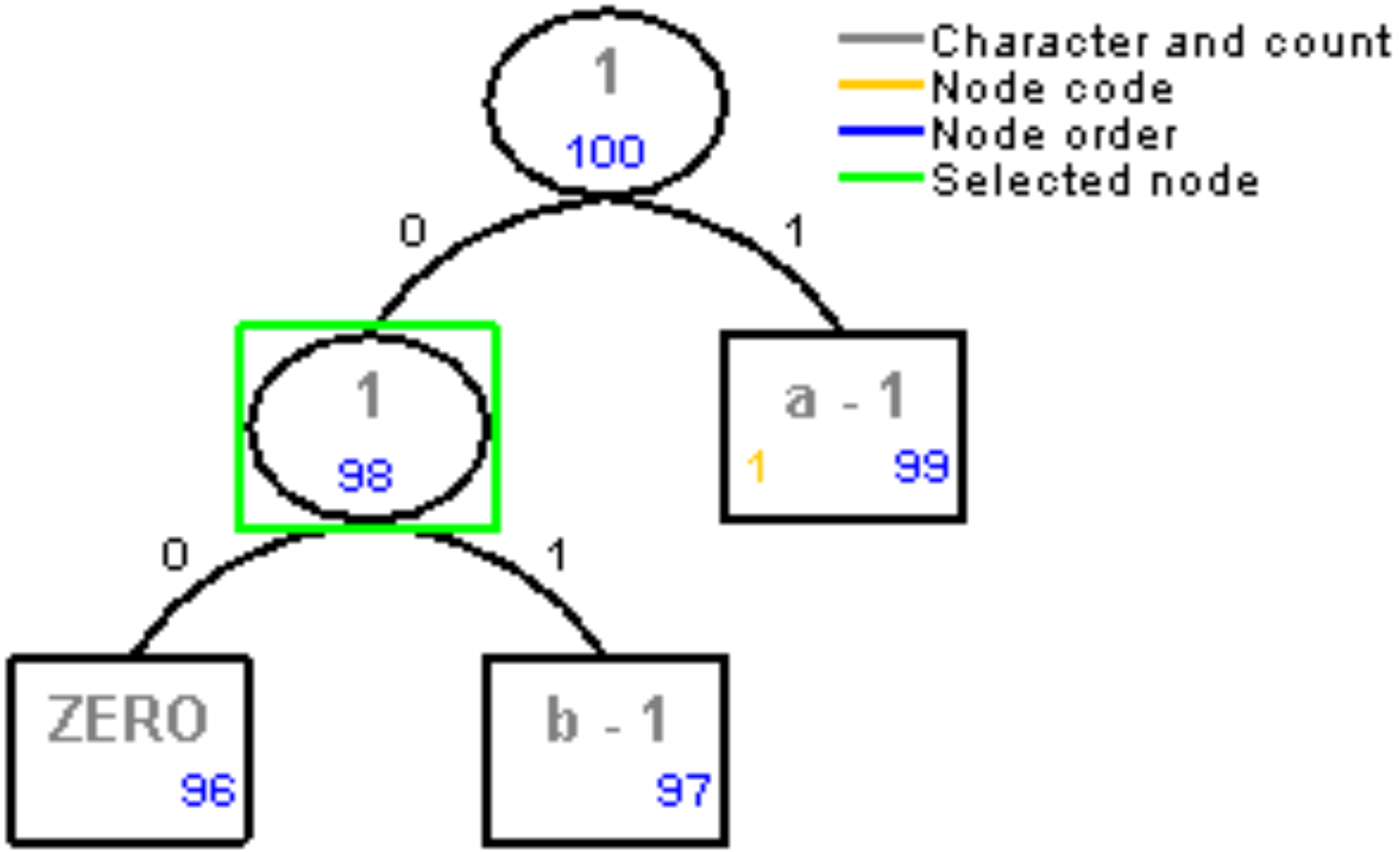
FGK Example - abba



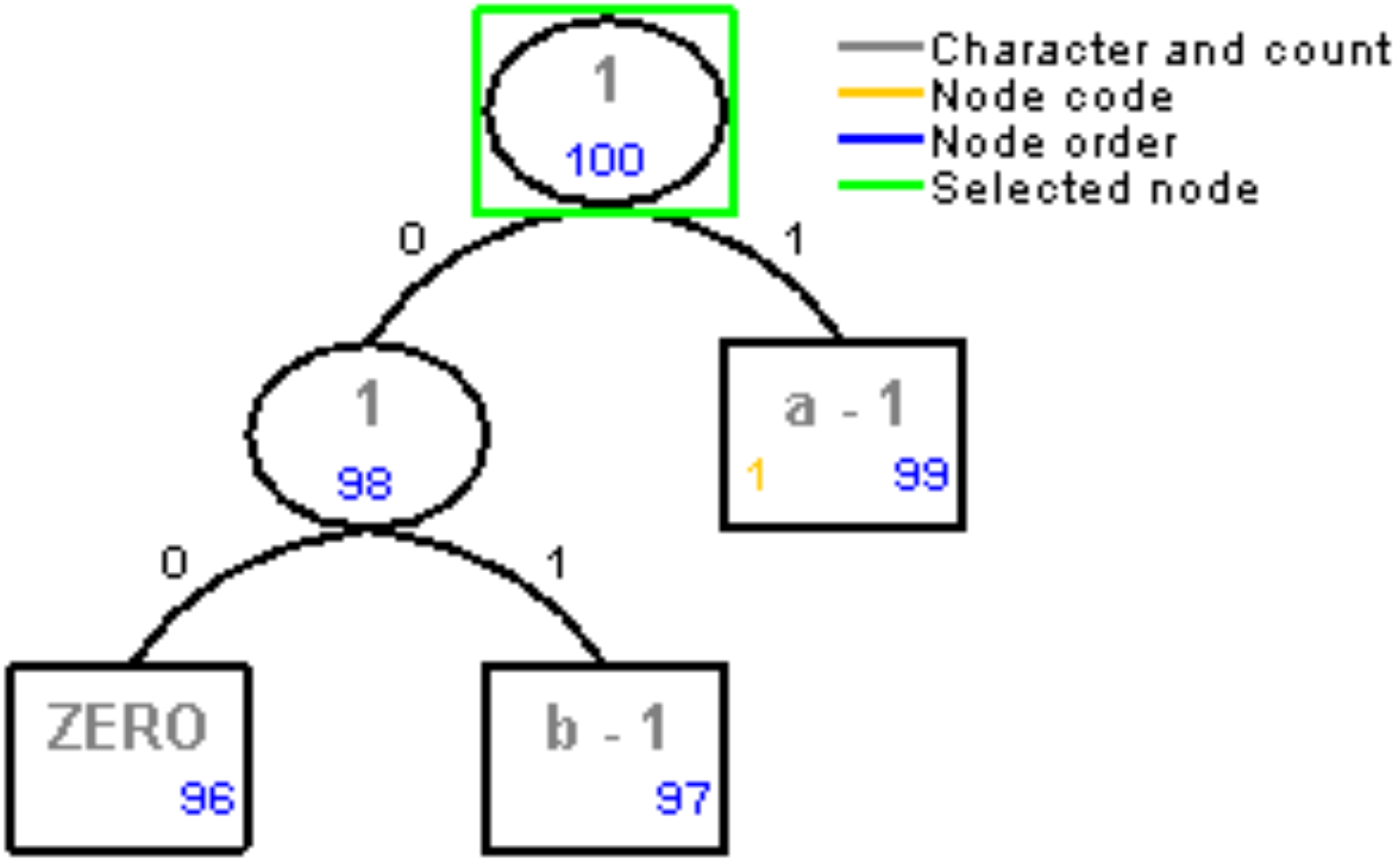
FGK Example - abba



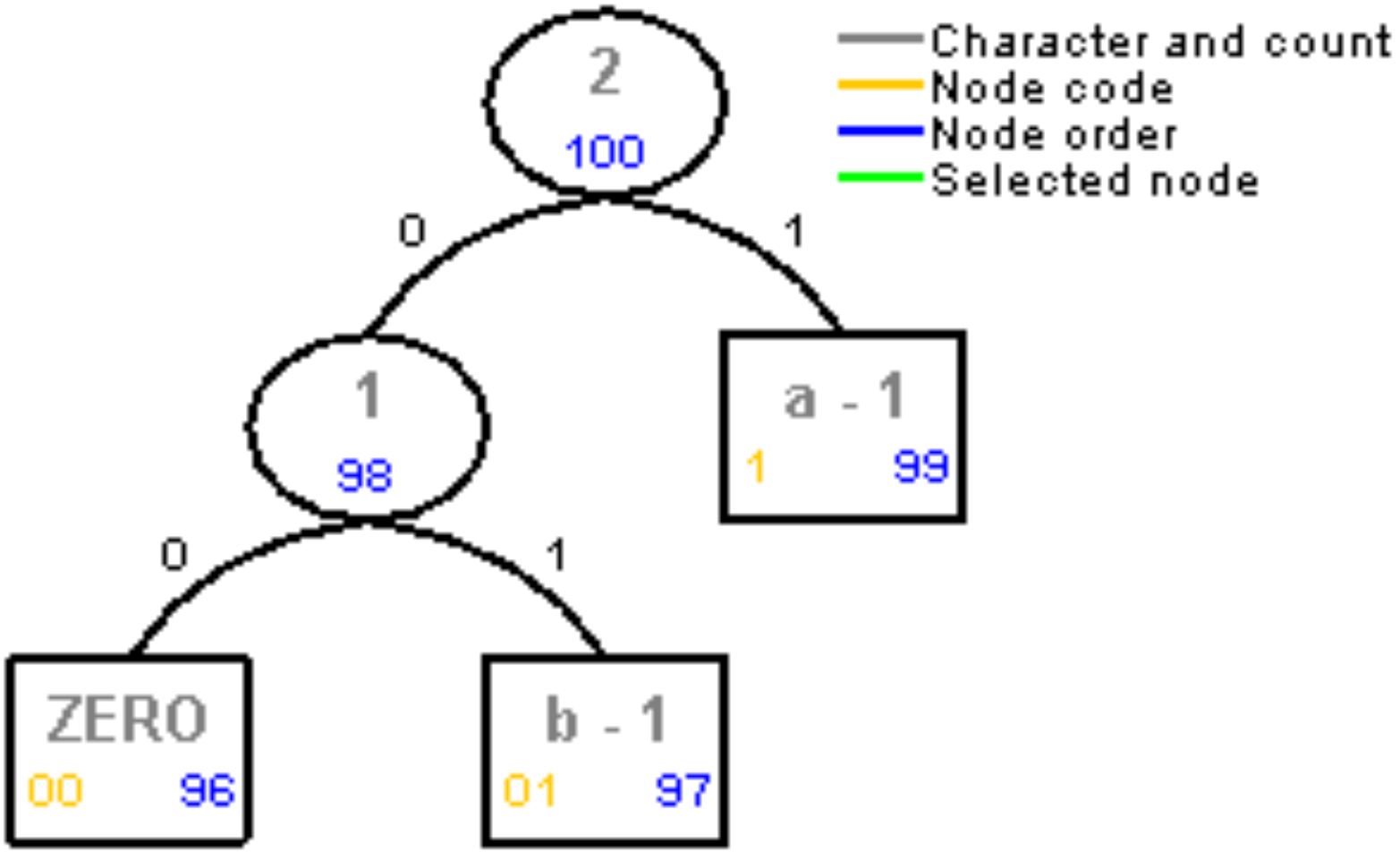
FGK Example - abba



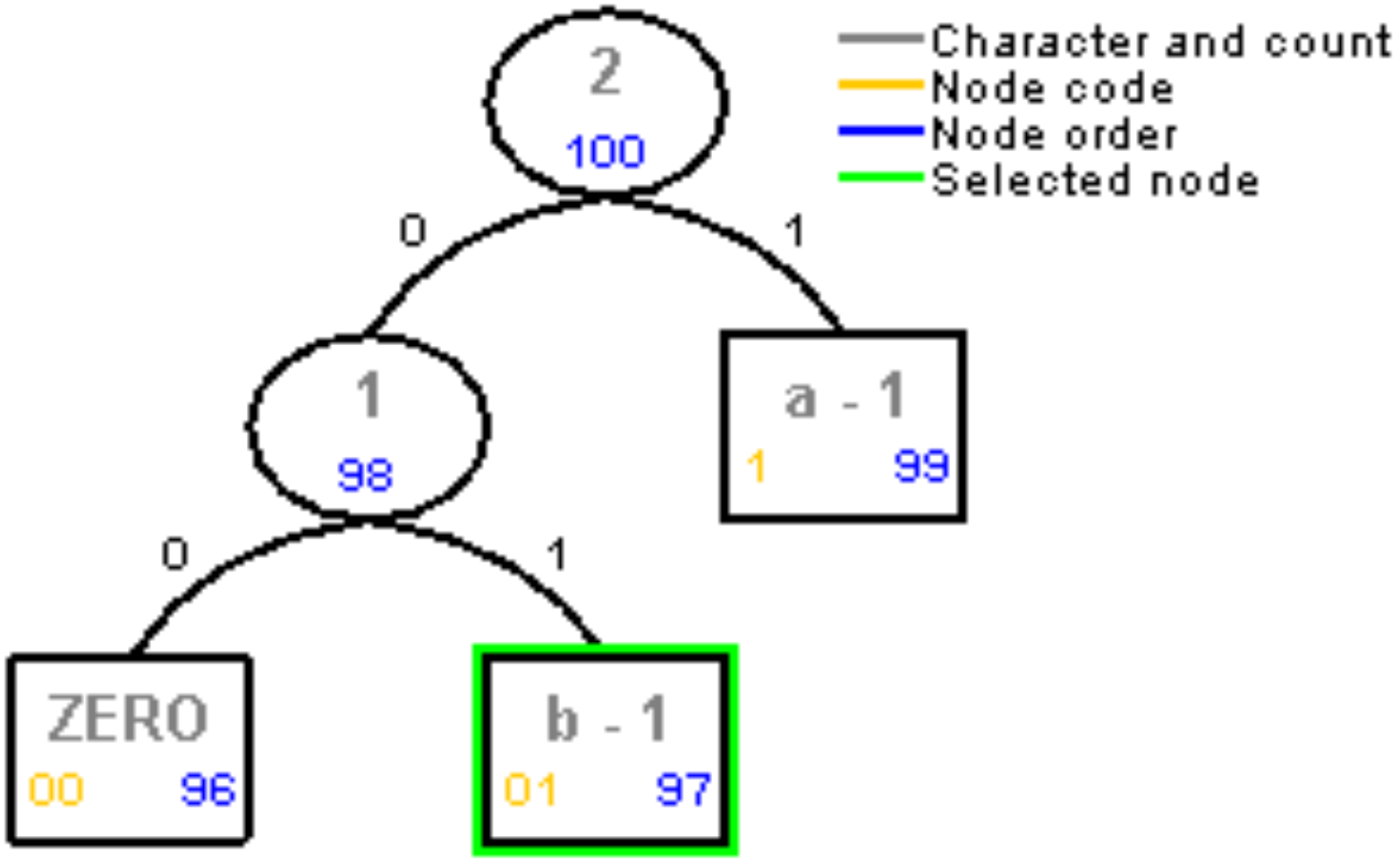
FGK Example - abba



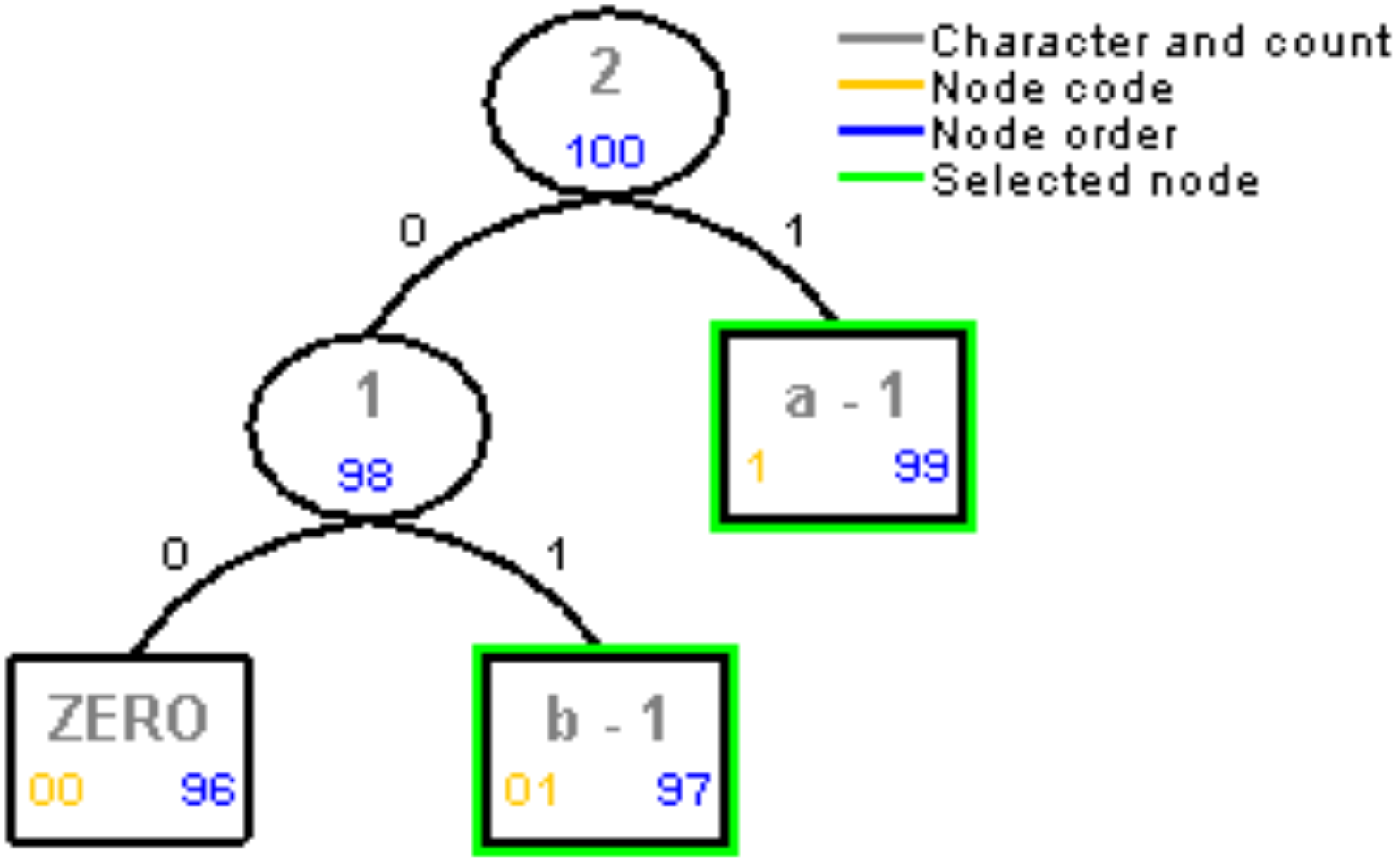
FGK Example - abba



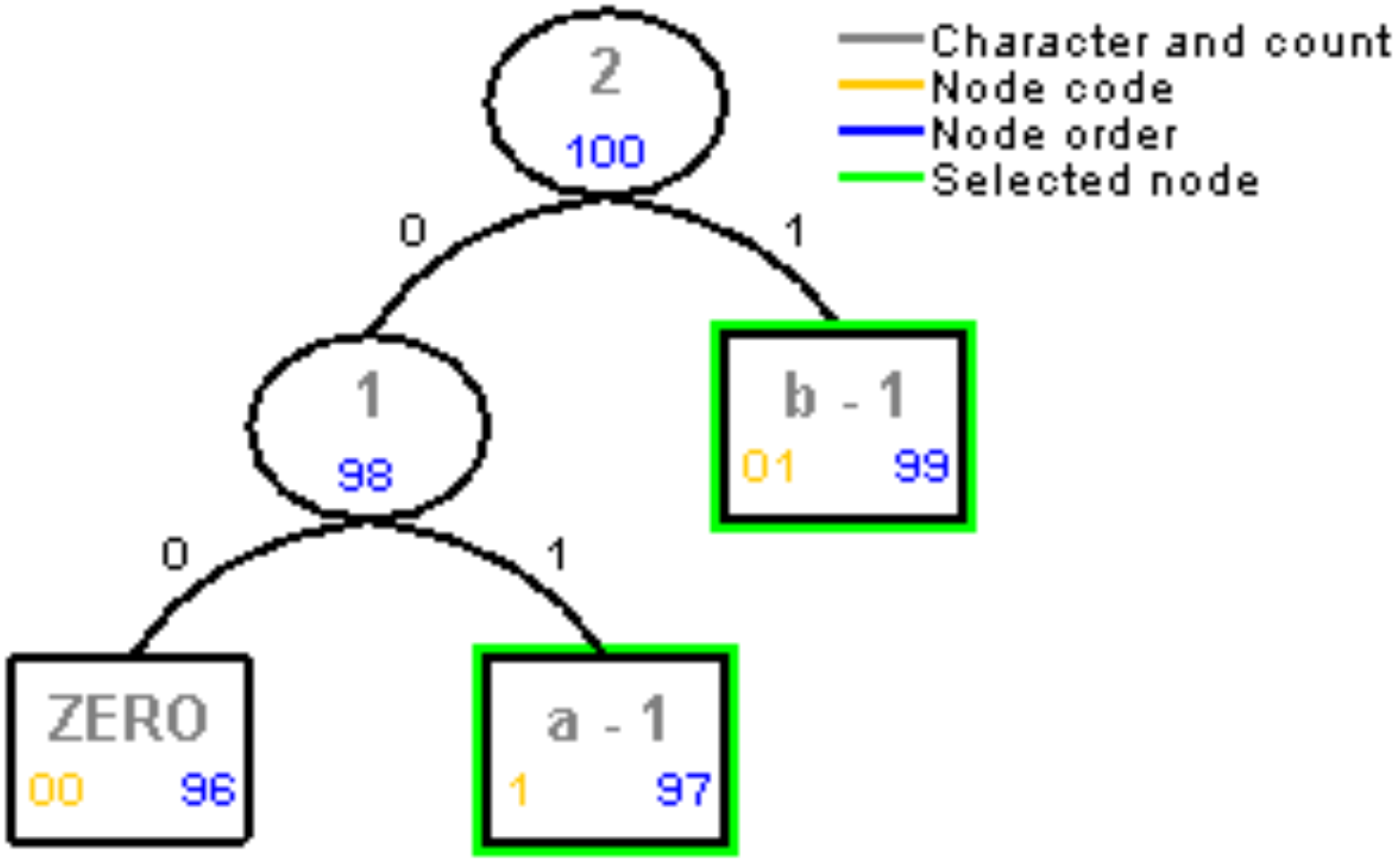
FGK Example - abba



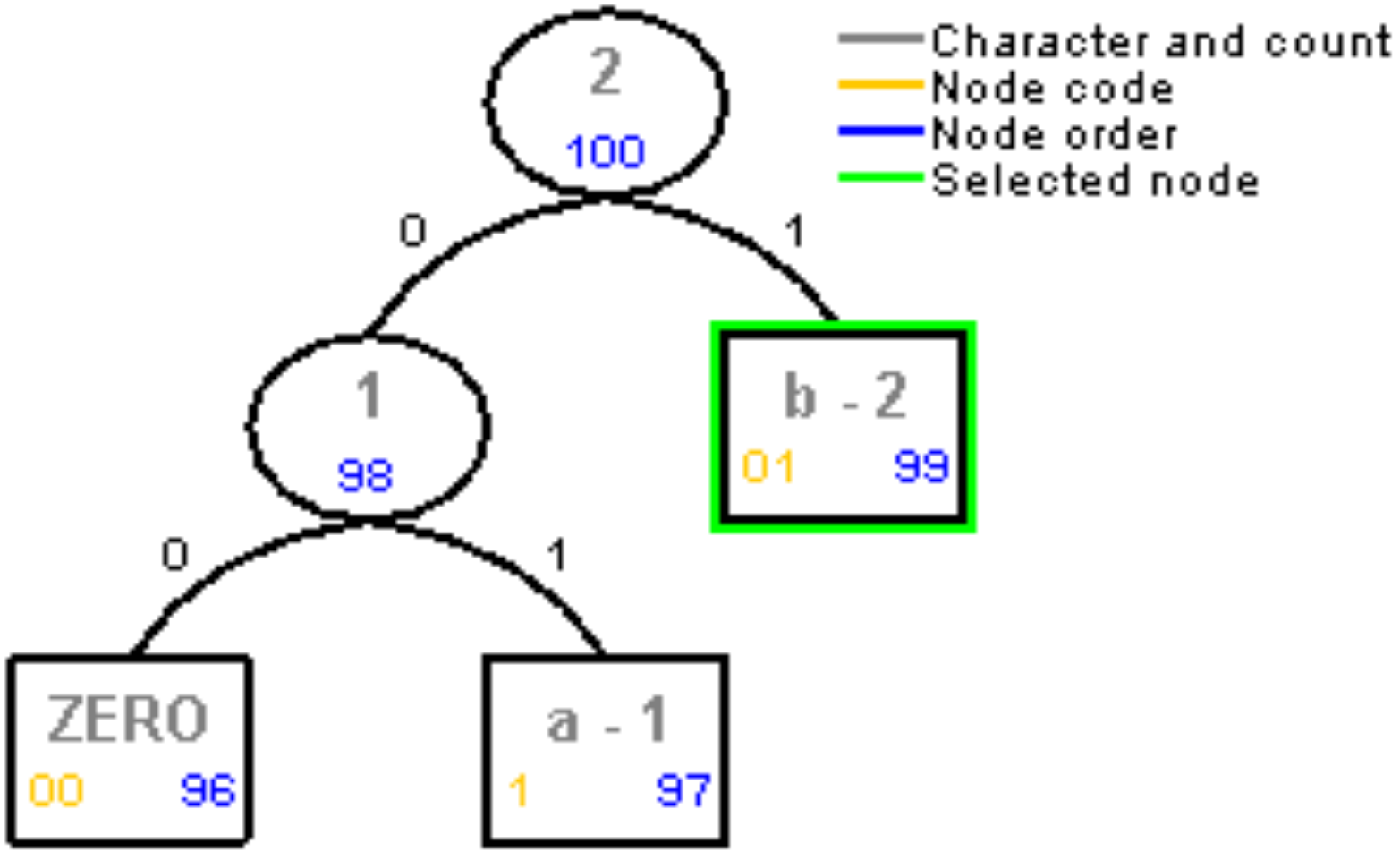
FGK Example - abba



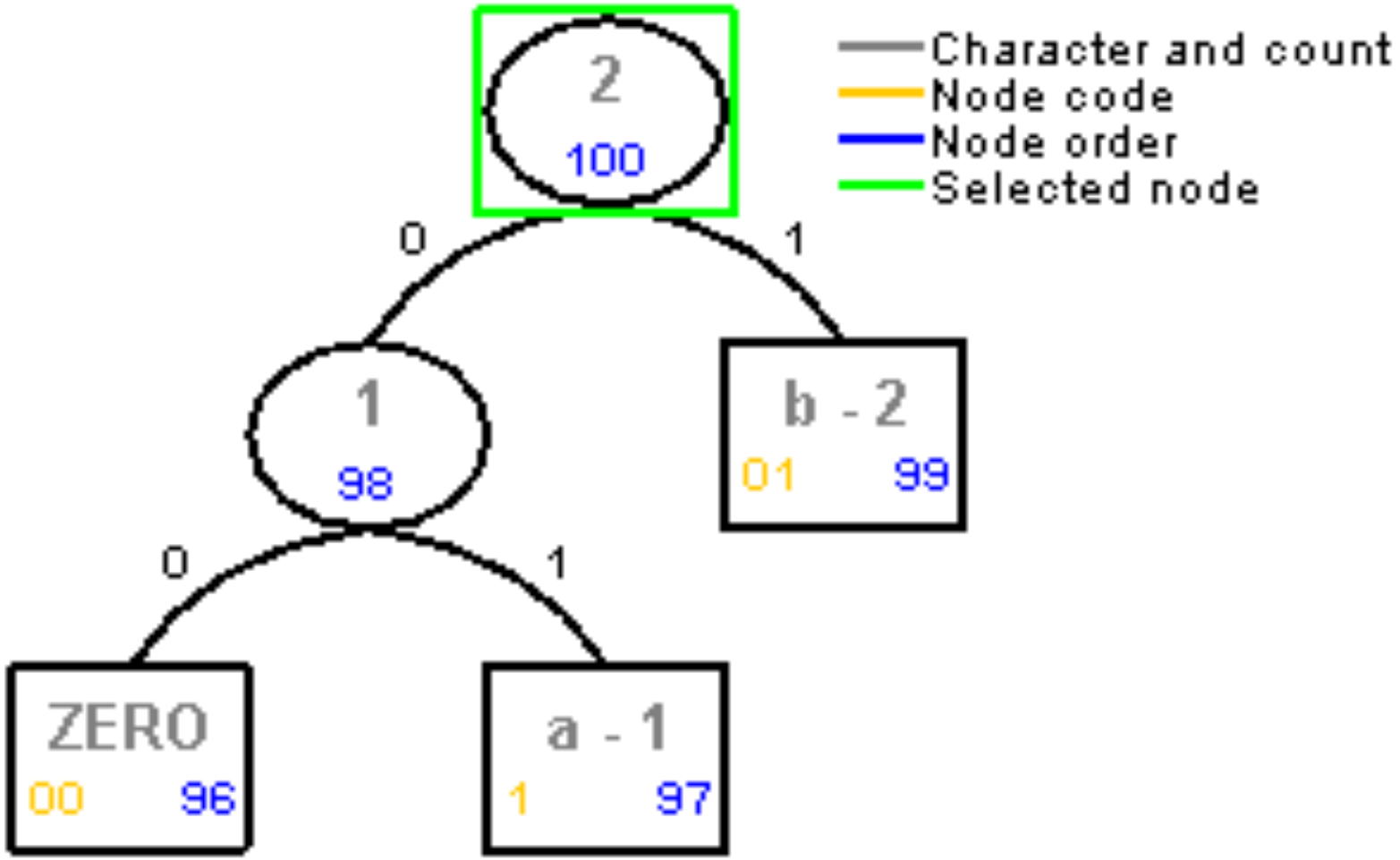
FGK Example - abba



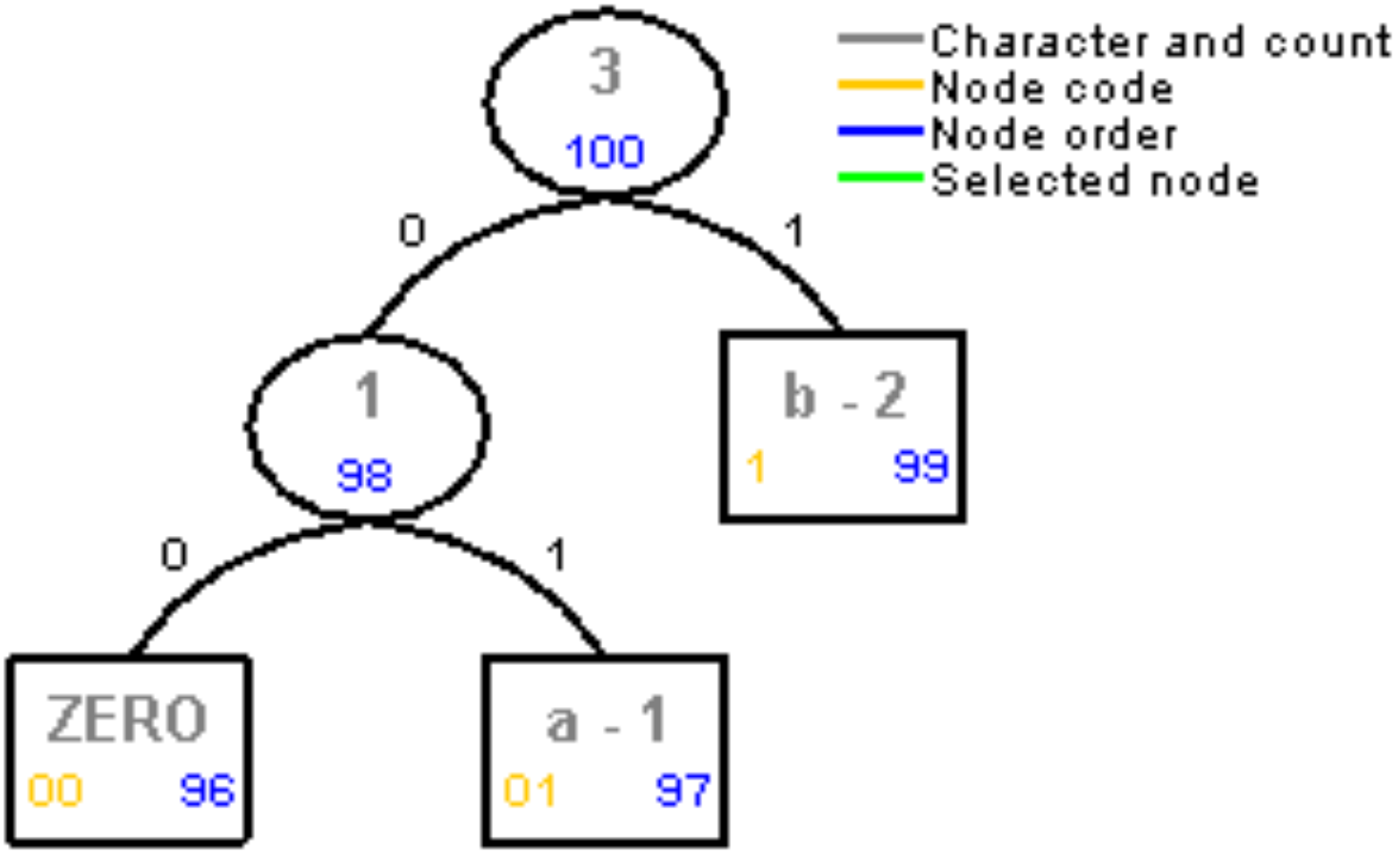
FGK Example - abba



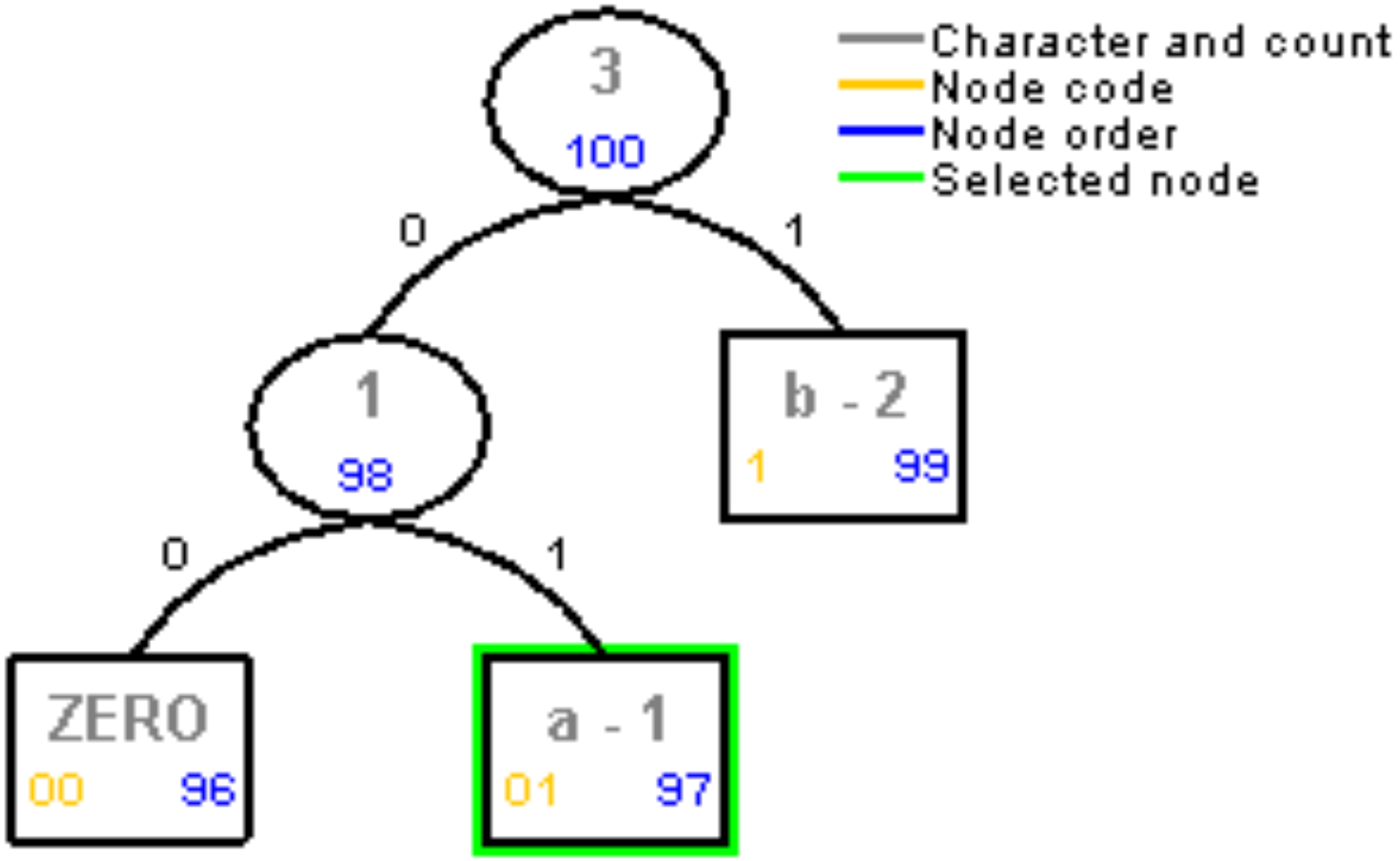
FGK Example - abba



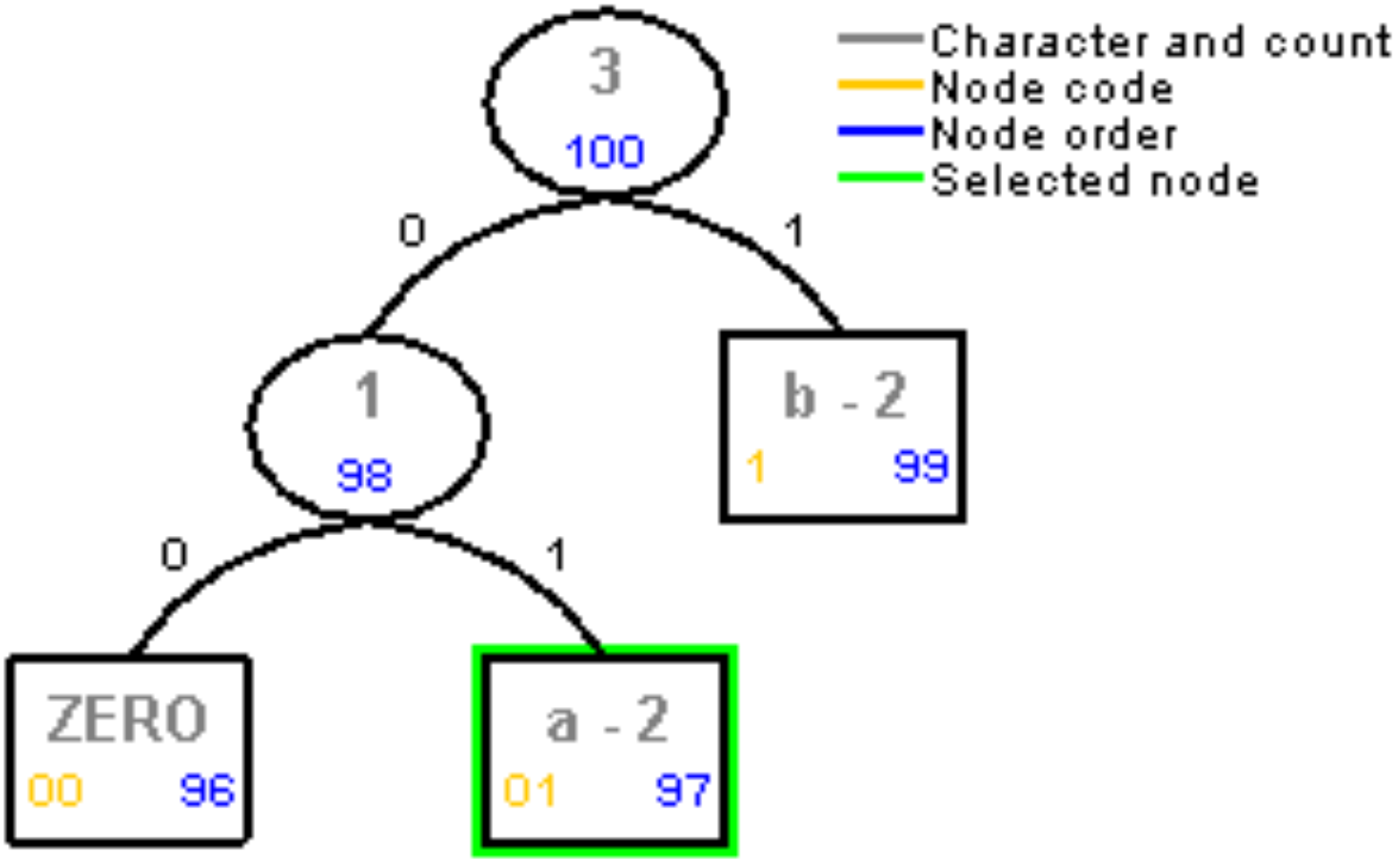
FGK Example - abba



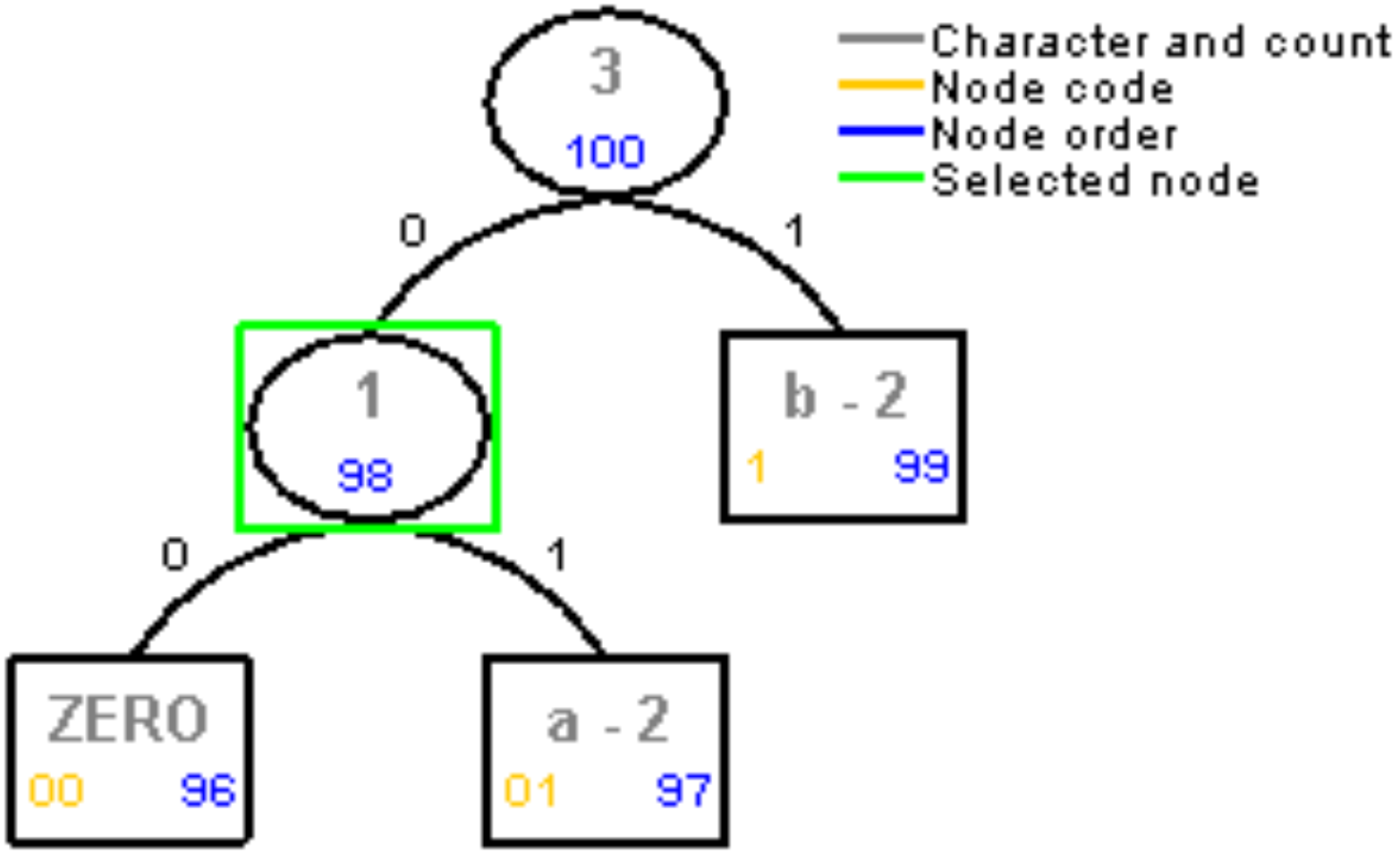
FGK Example - abba



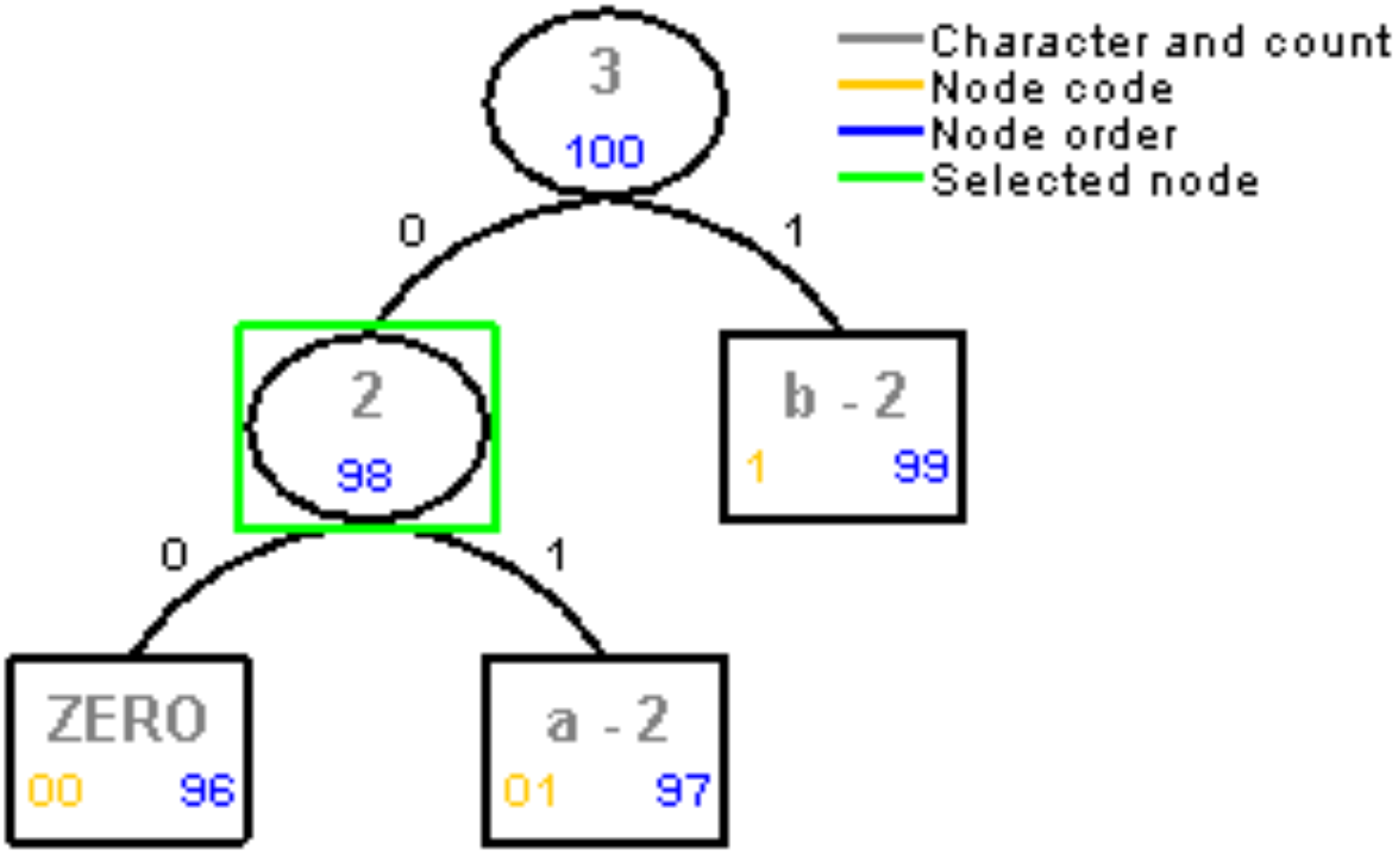
FGK Example - abba



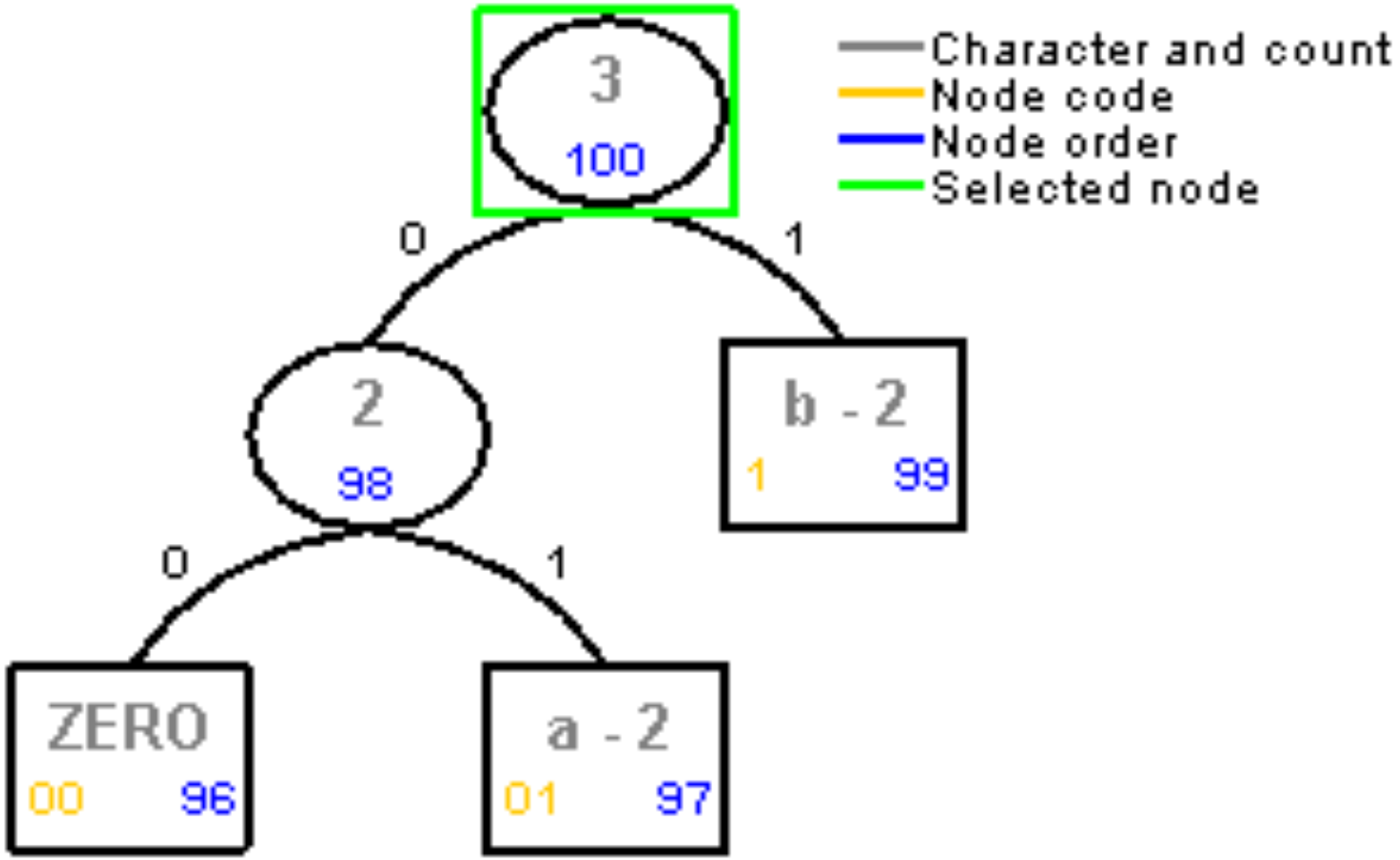
FGK Example - abba



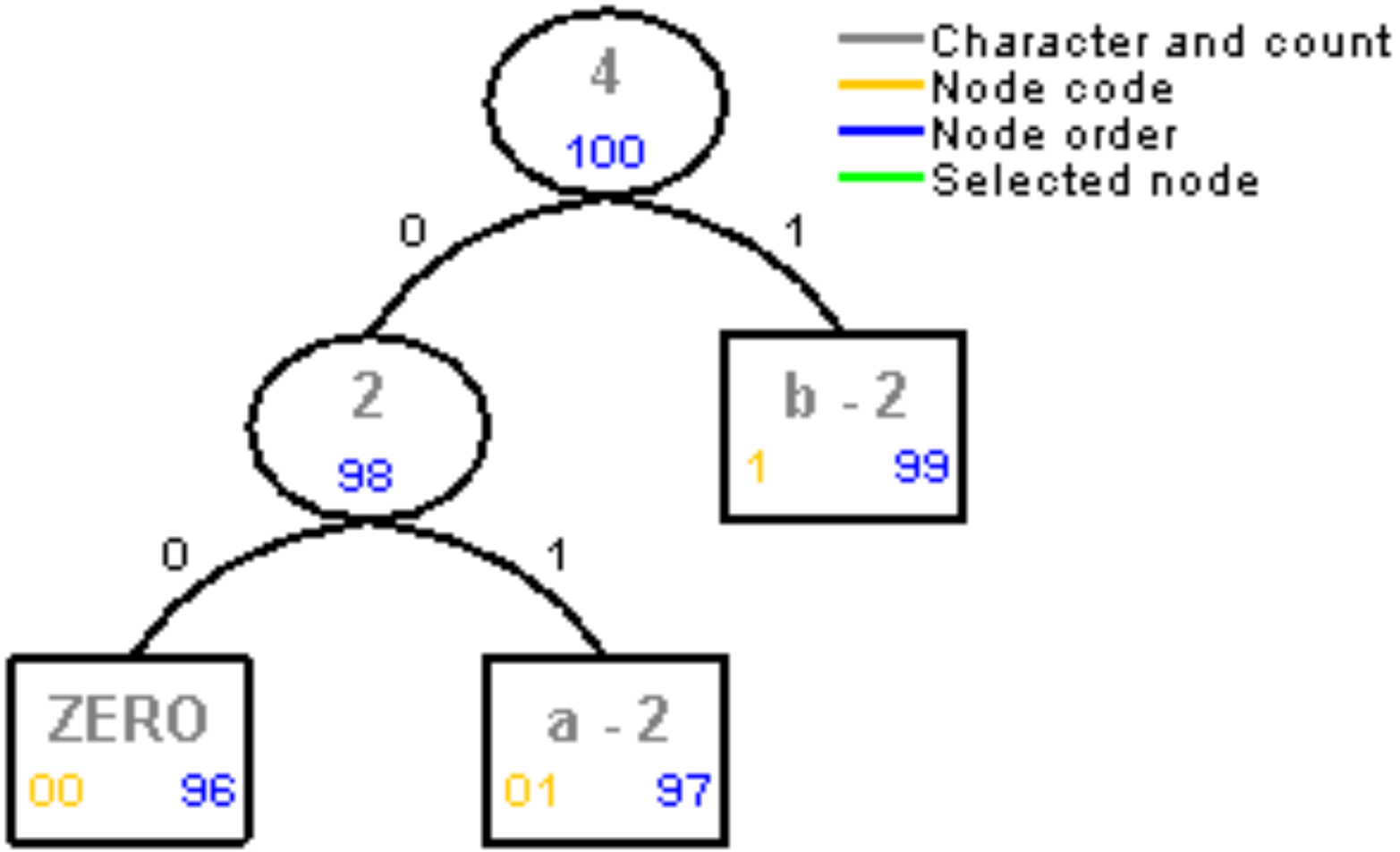
FGK Example - abba



FGK Example - abba



FGK Example - abba

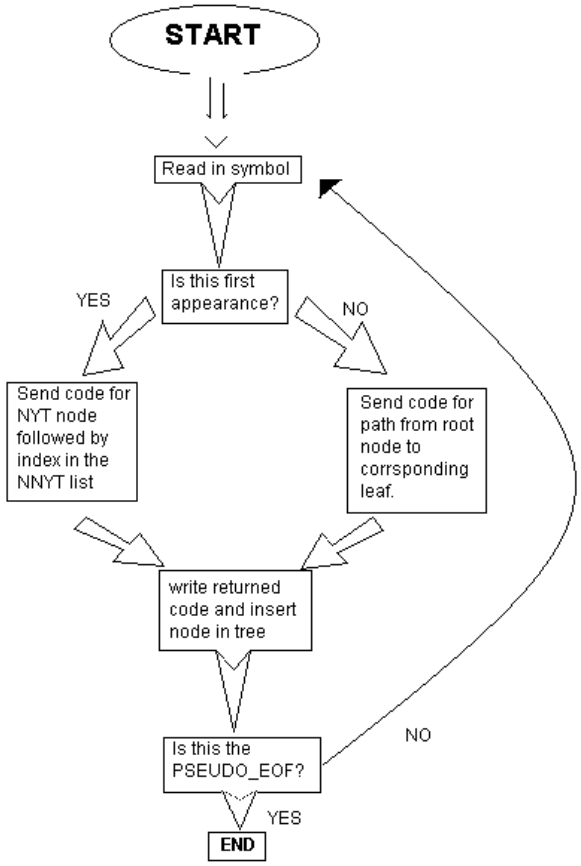


Vitter's algorithm

- quite similar to FGK
- uses another way to update coding tree
- results into more balanced tree.

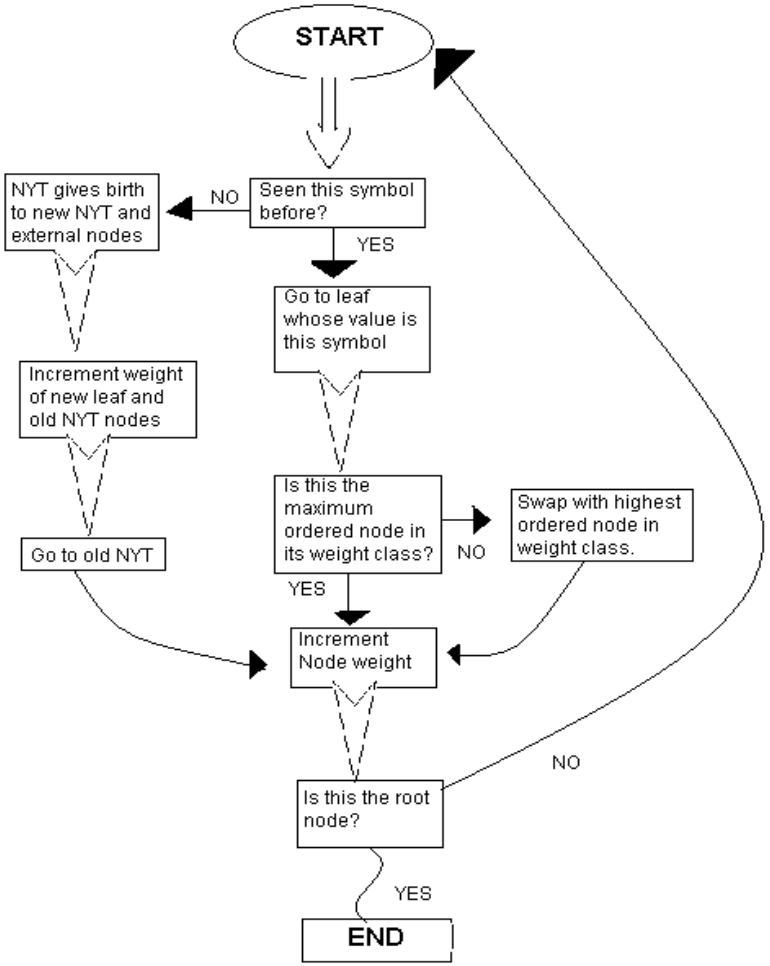
Vitter's algorithm

encoding
procedure



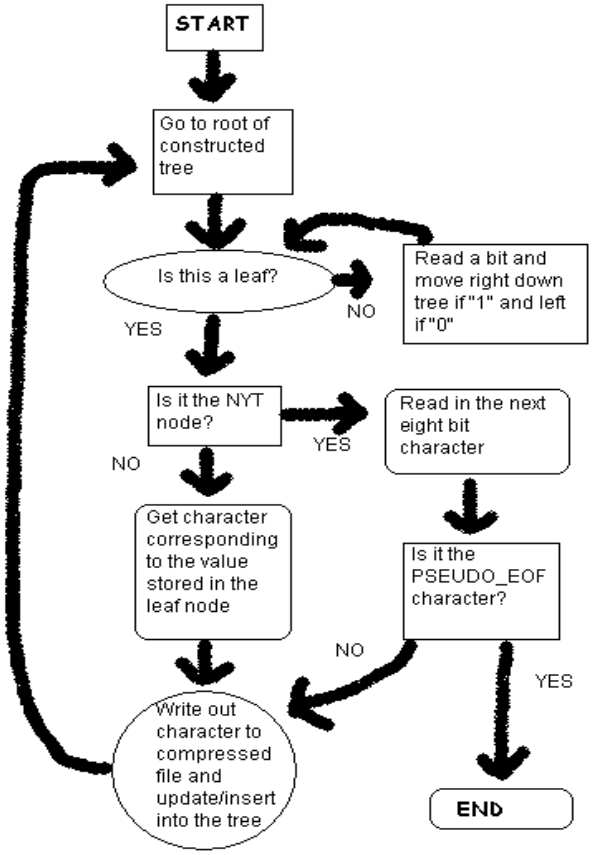
Vitter's algorithm

tree manipulation process

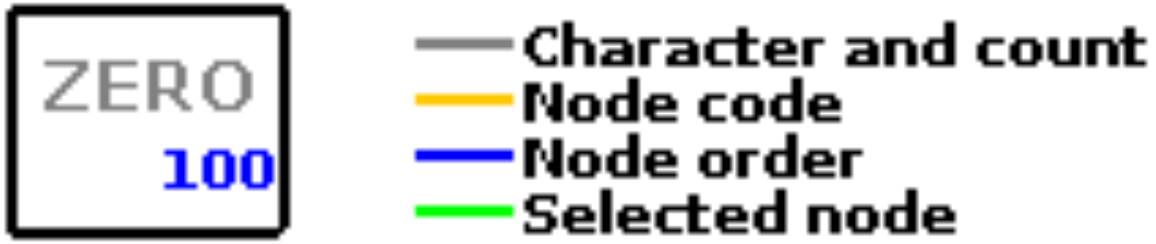


Vitter's algorithm

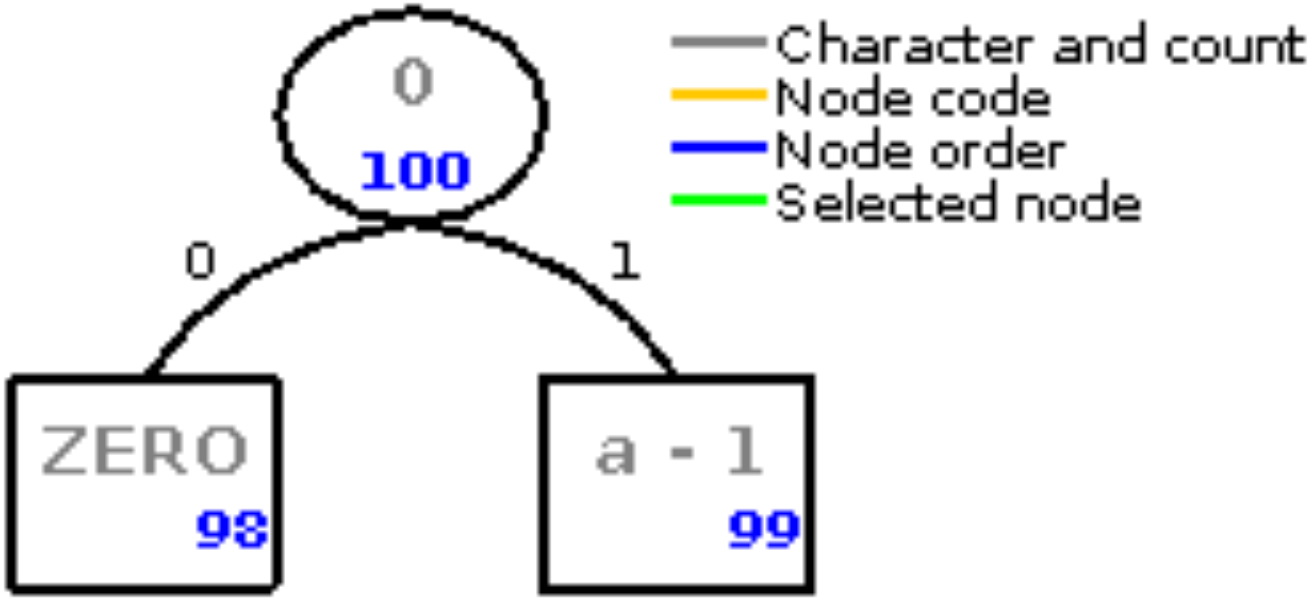
decoding procedure



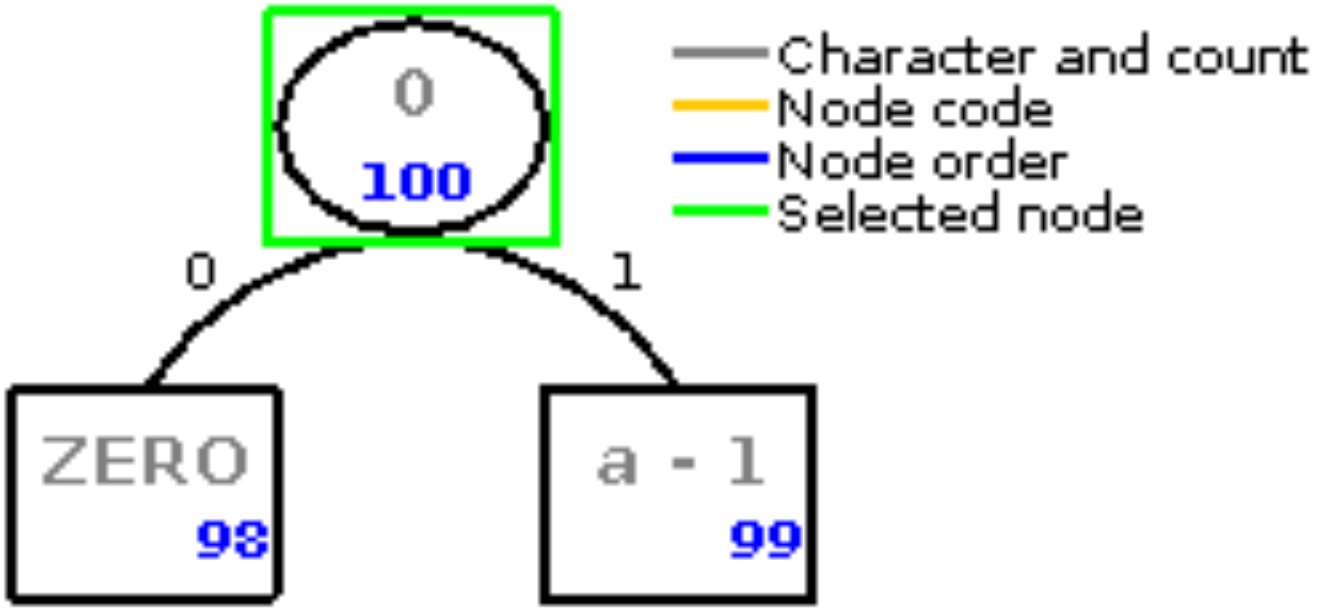
Vitter example: abba



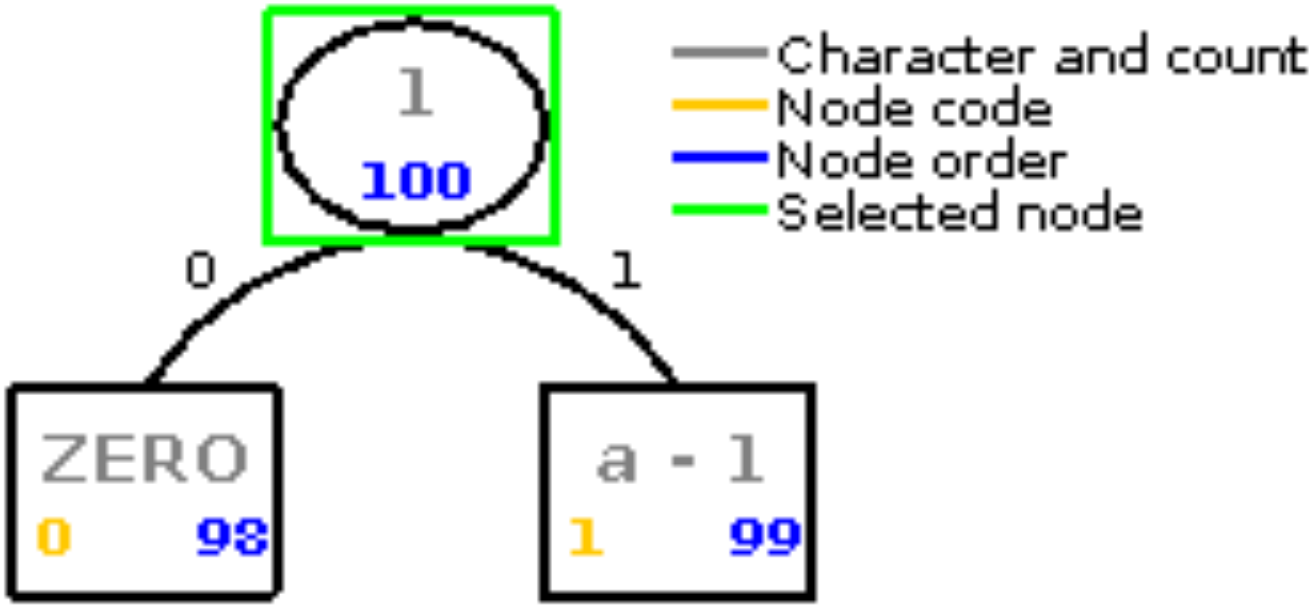
Vitter example: abba



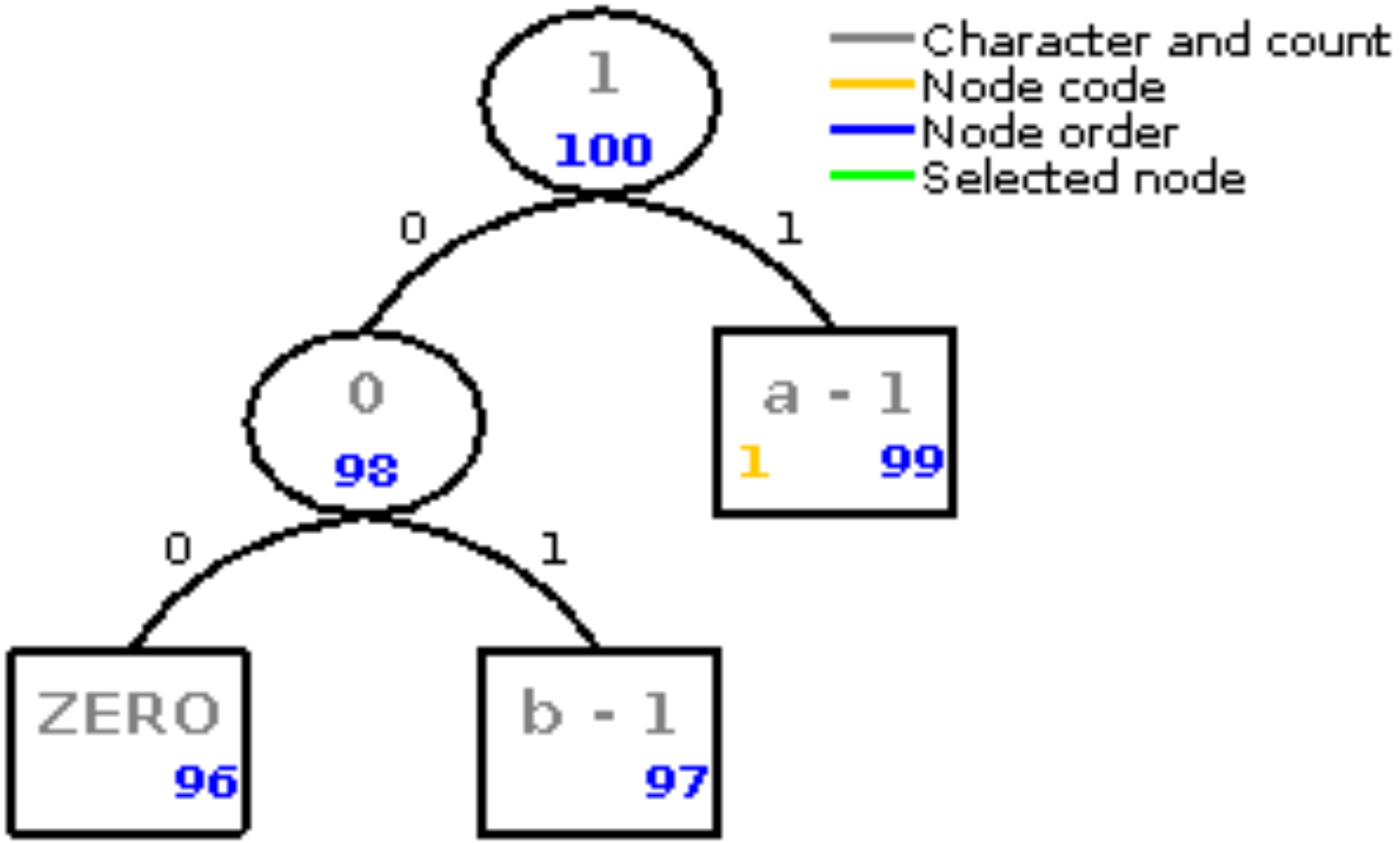
Vitter example: abba



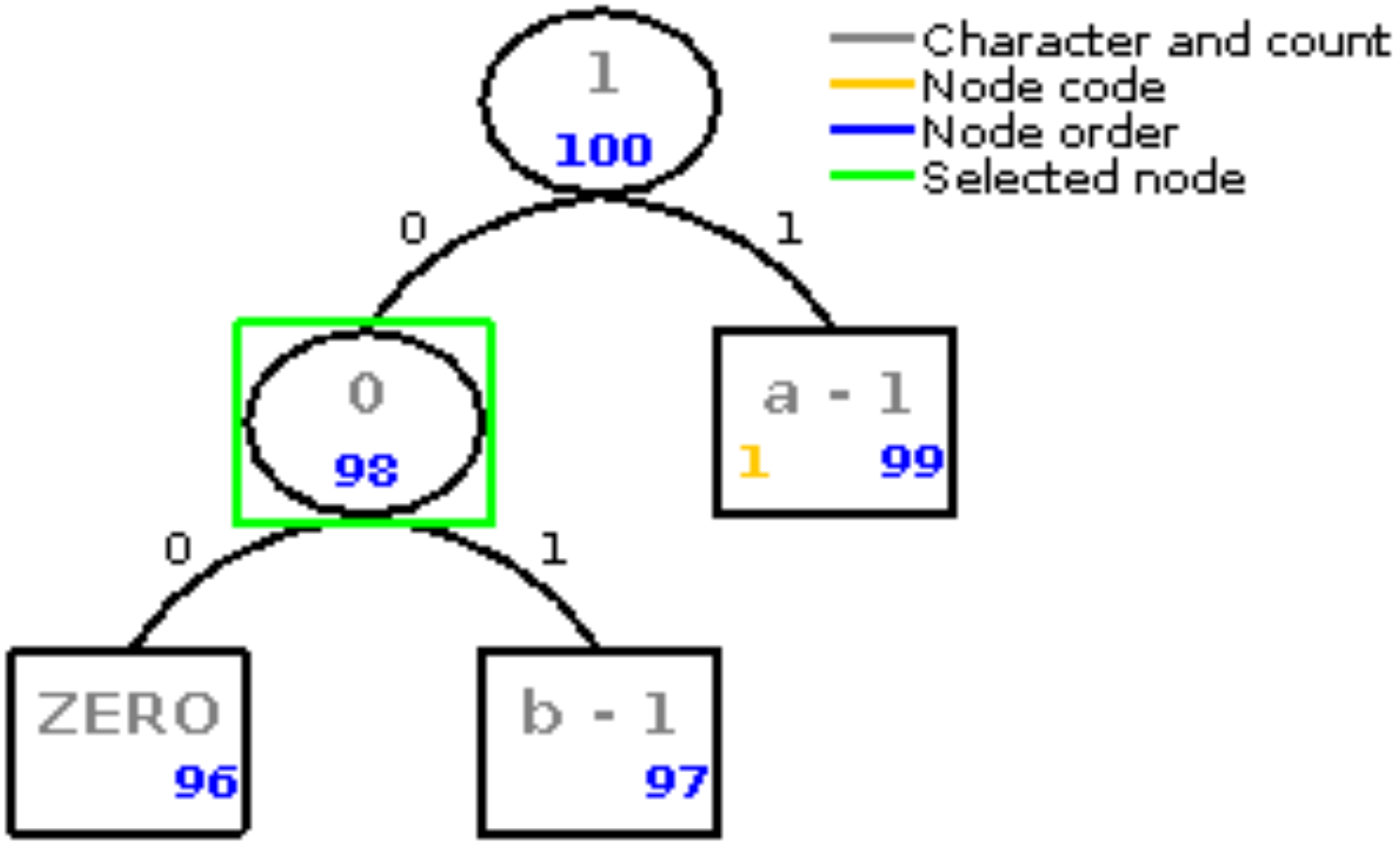
Vitter example: abba



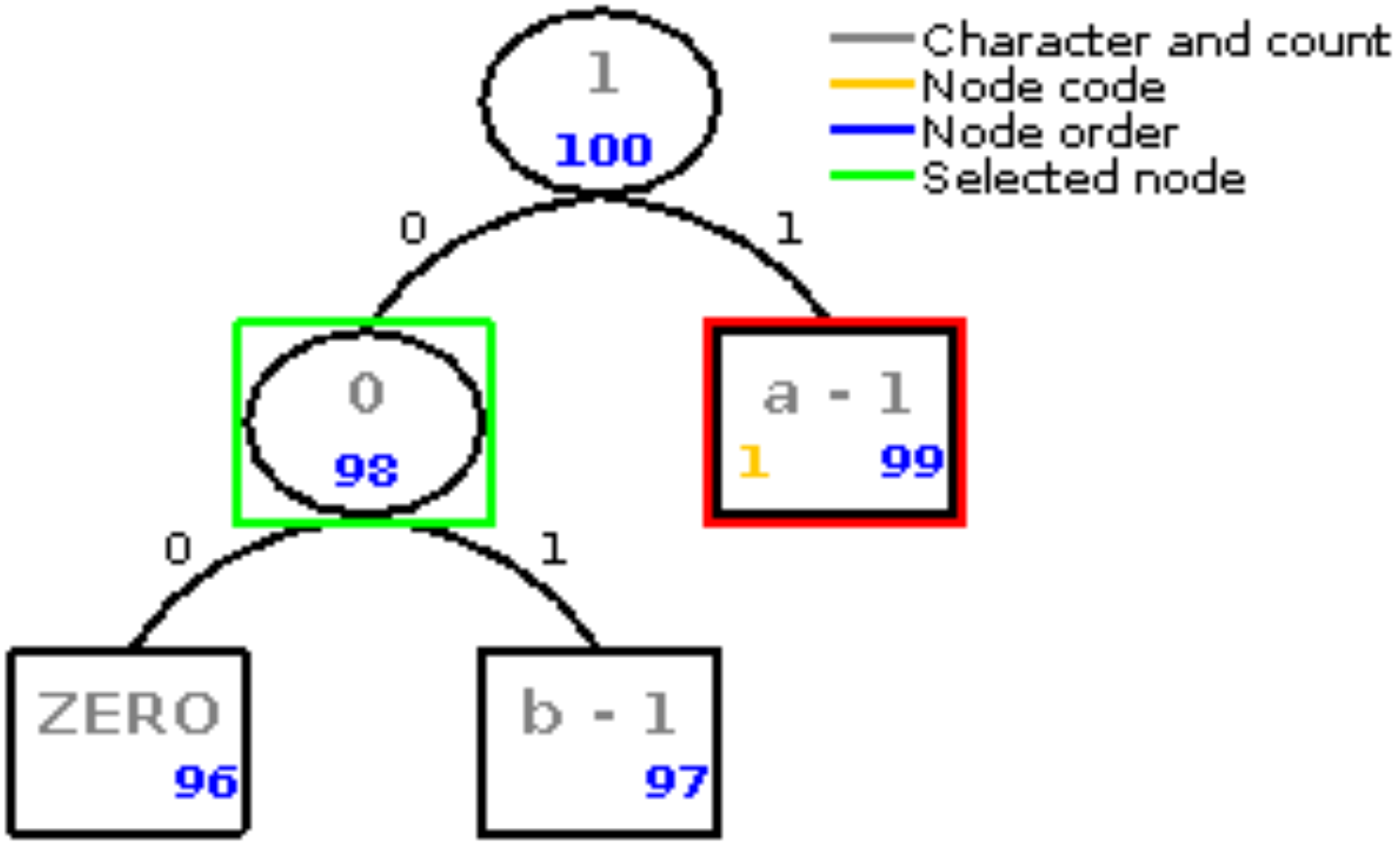
Vitter example: abba



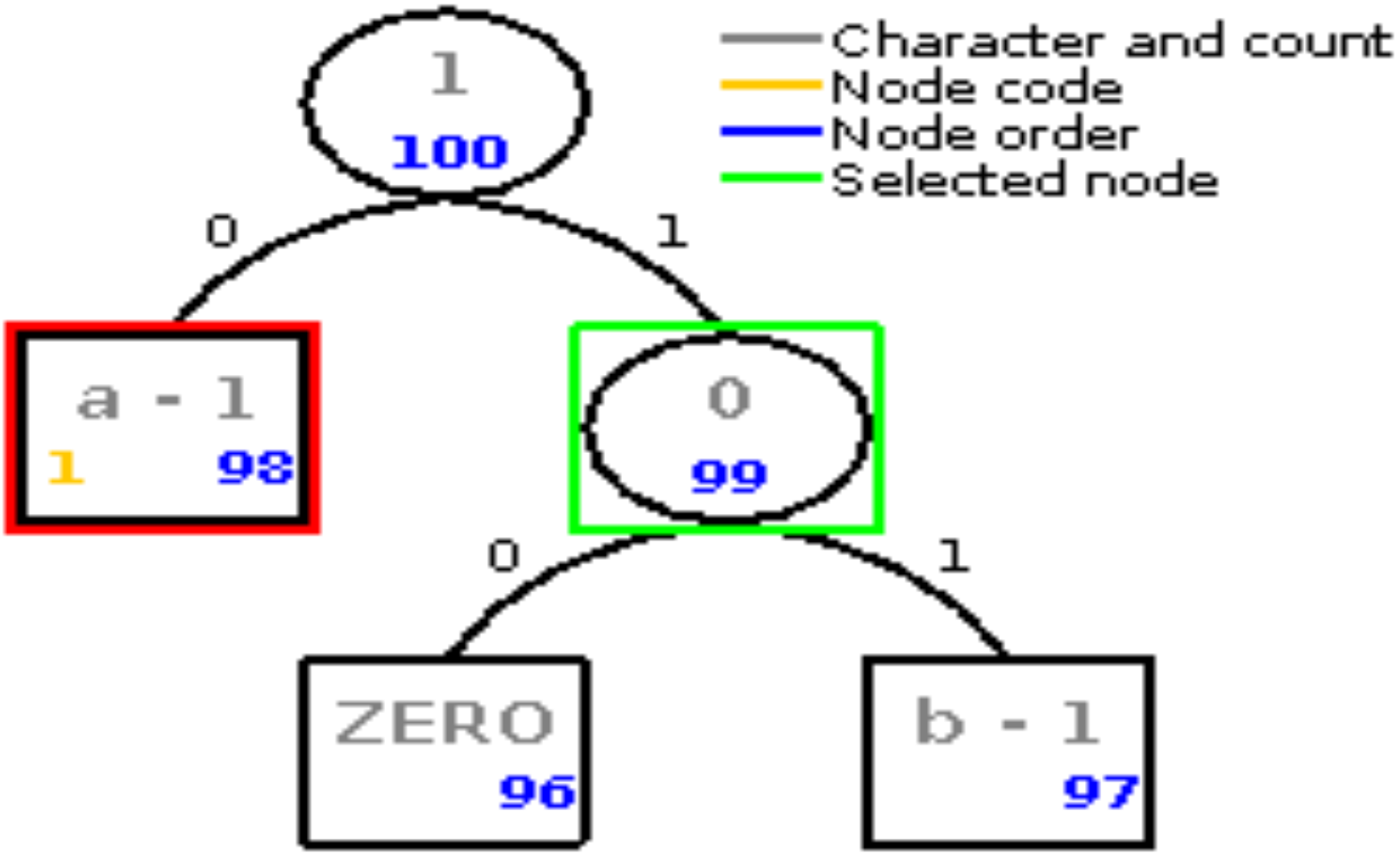
Vitter example: abba



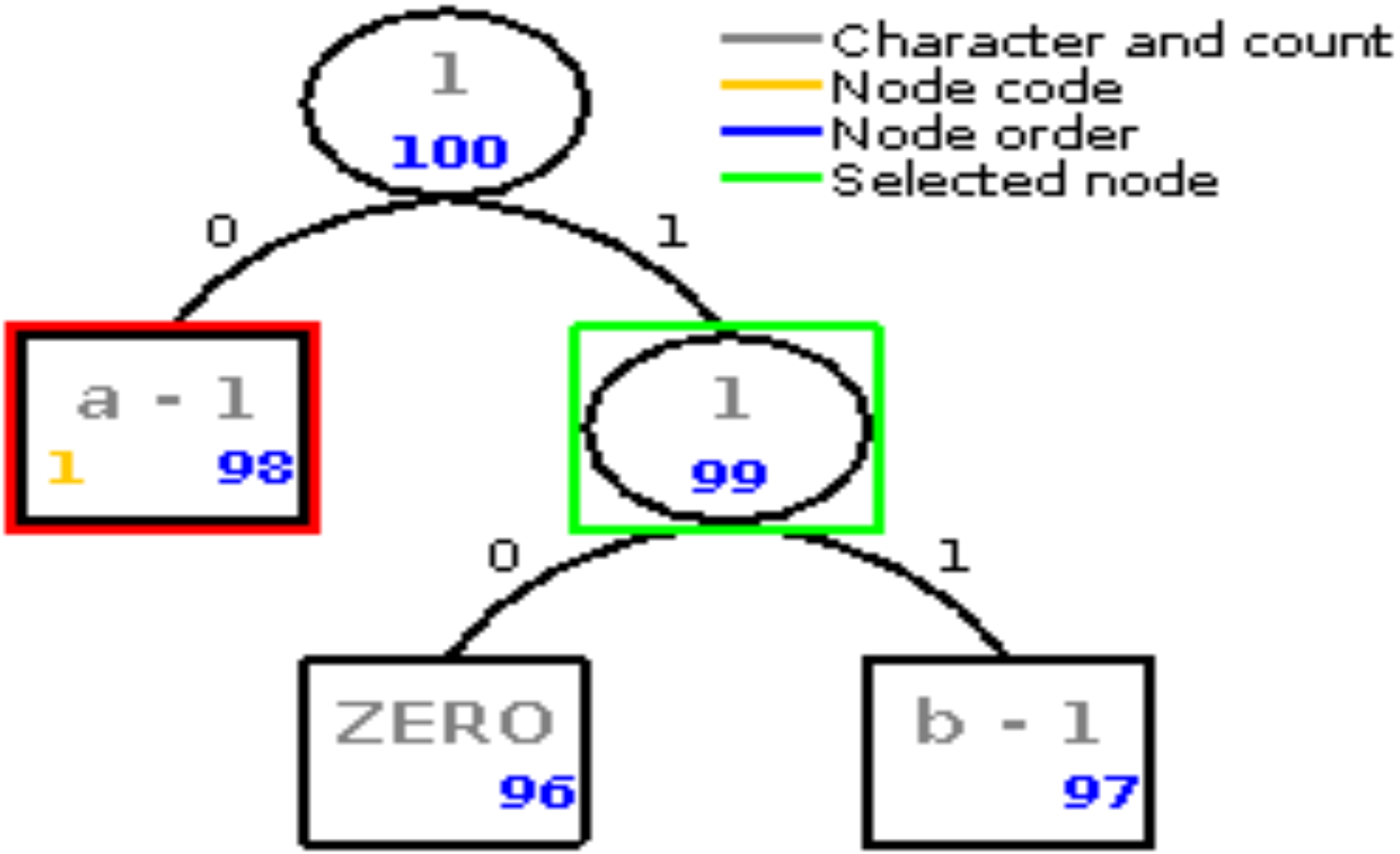
Vitter example: abba



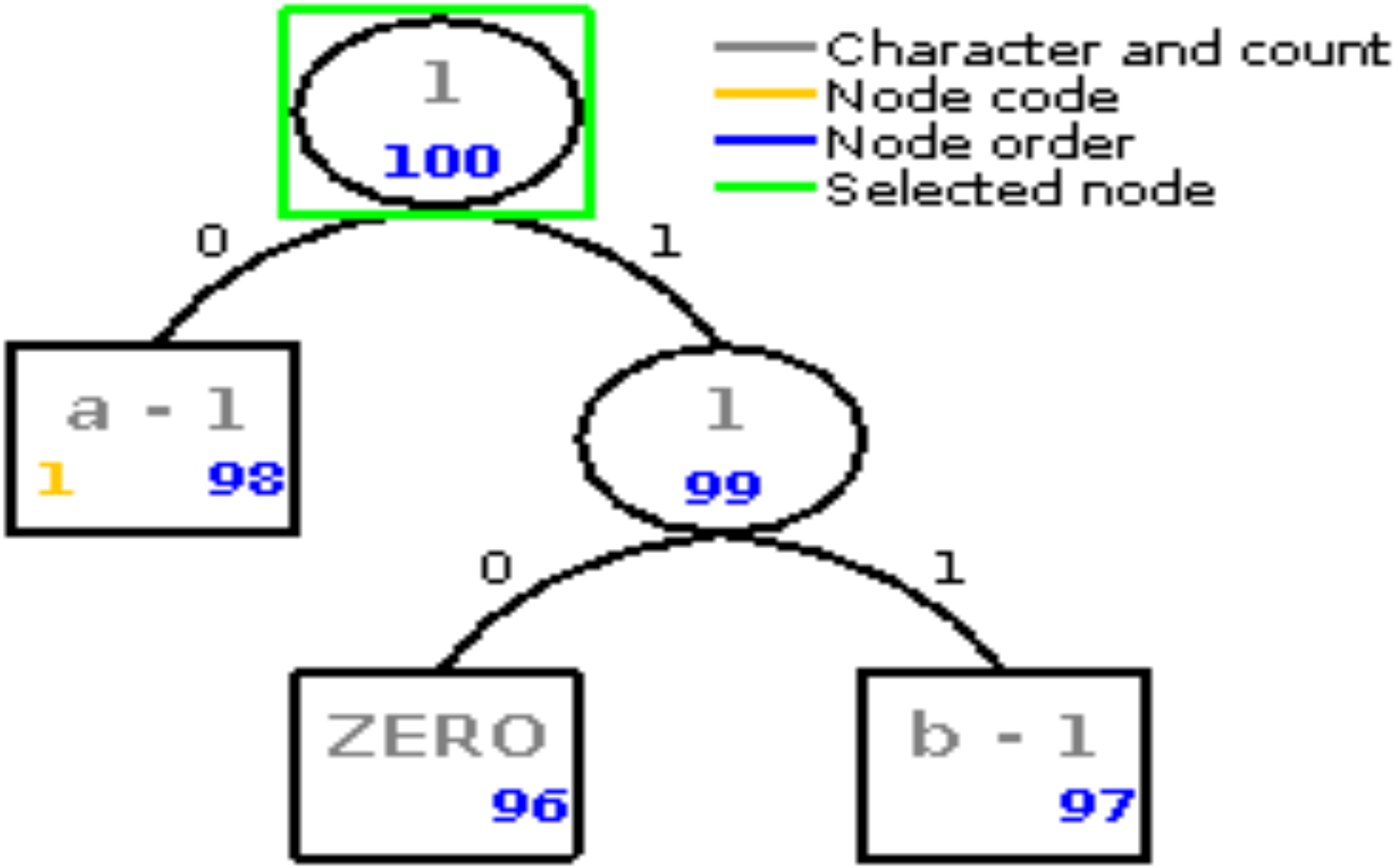
Vitter example: abba



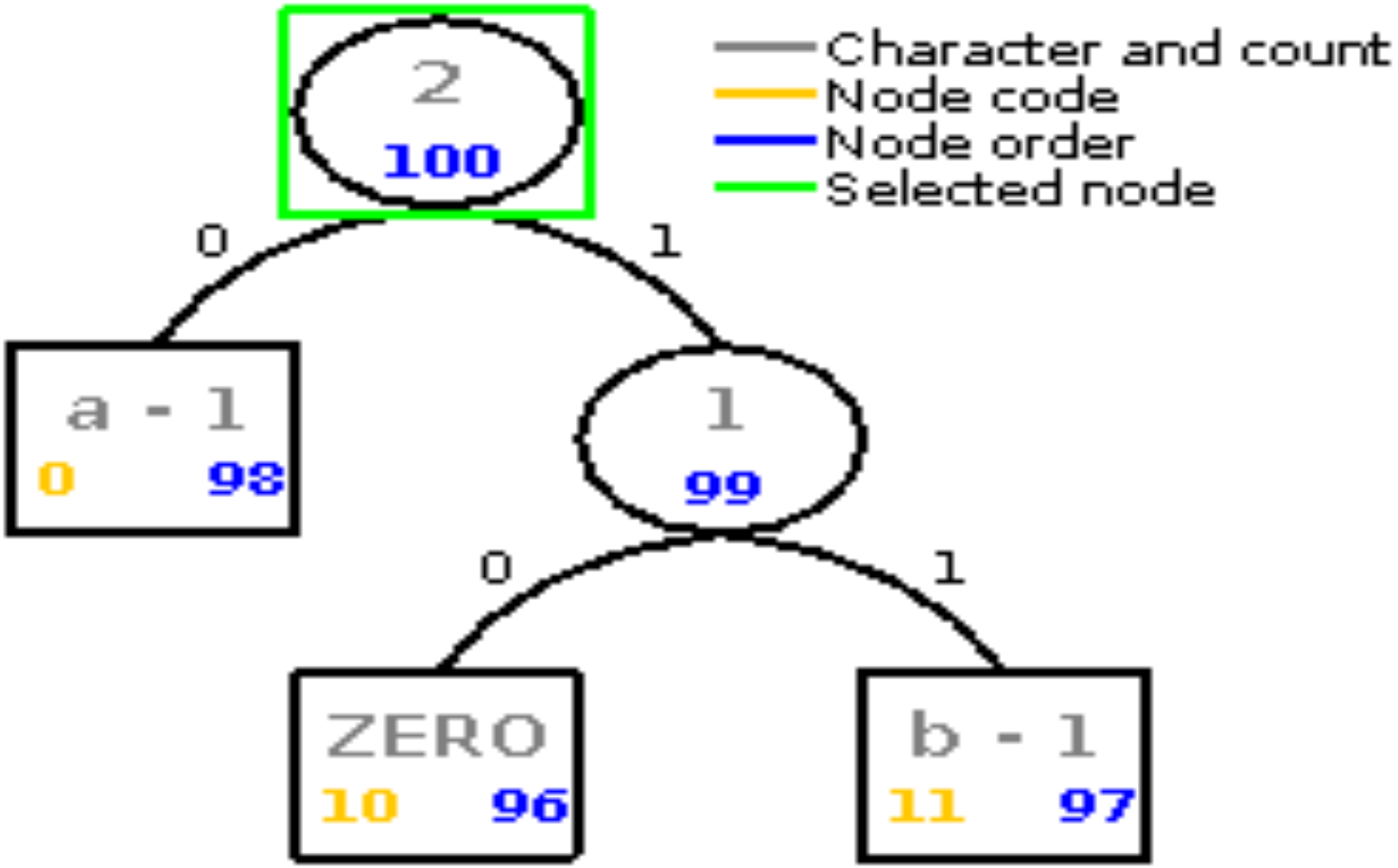
Vitter example: abba



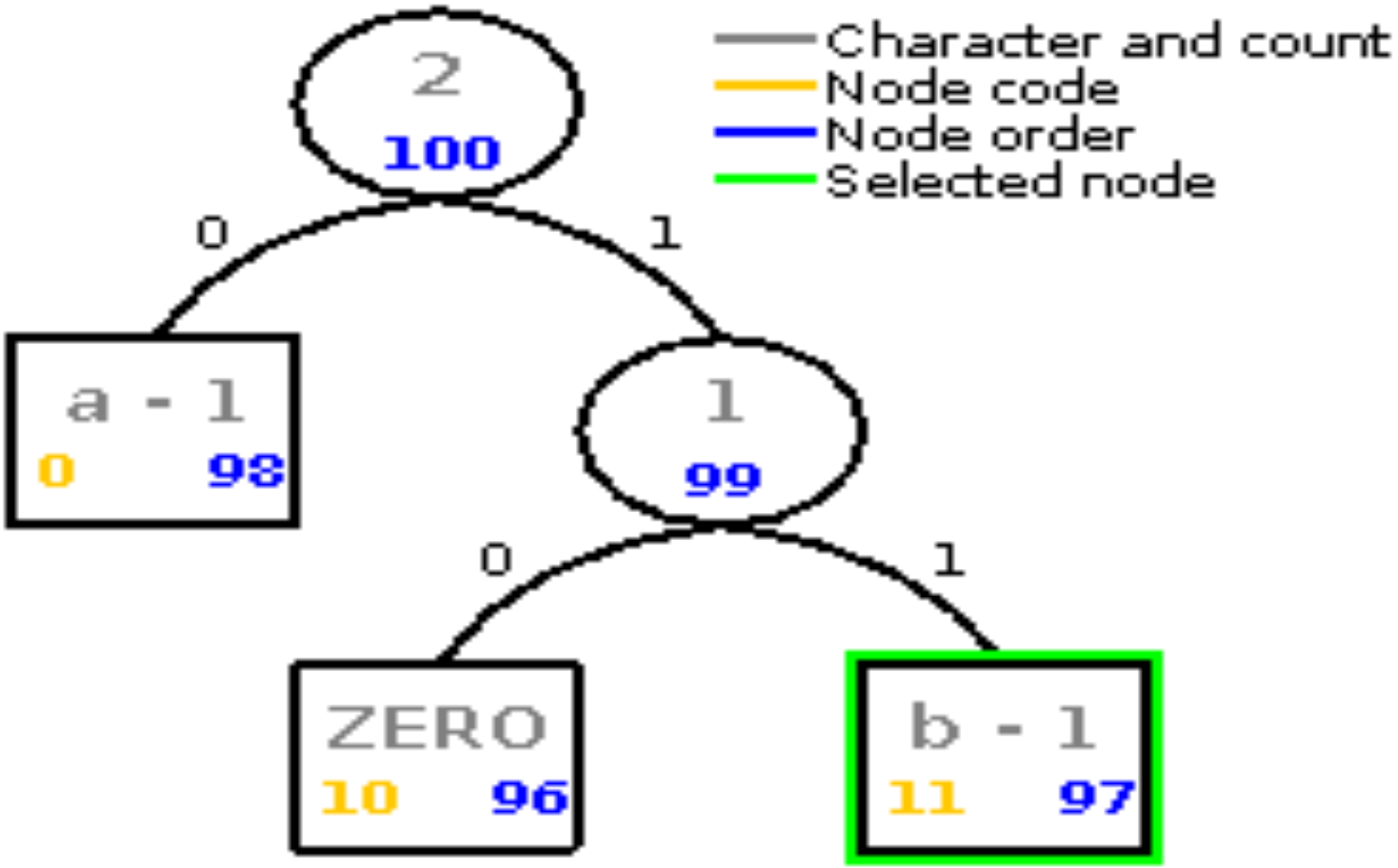
Vitter example: abba



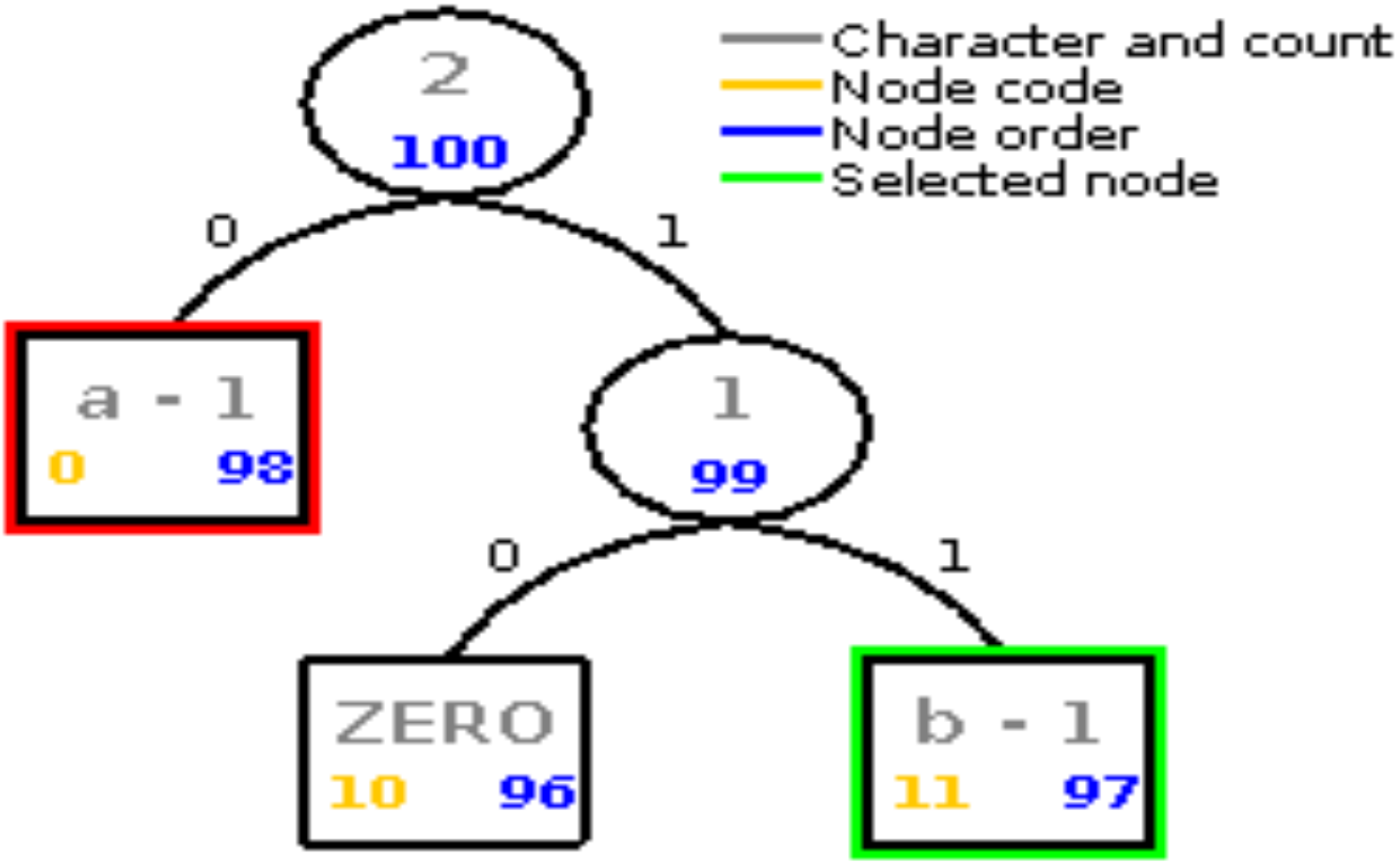
Vitter example: abba



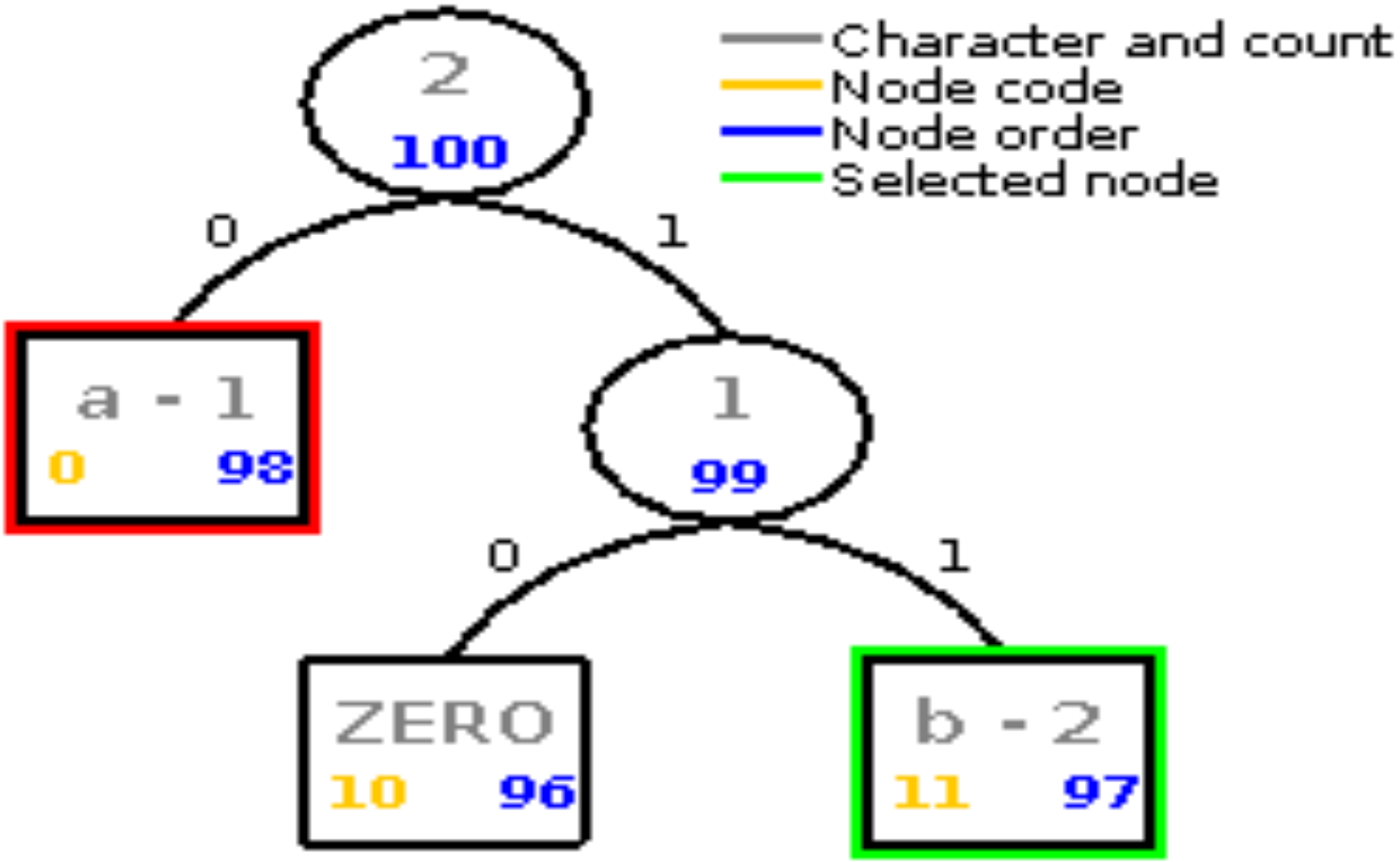
Vitter example: abba



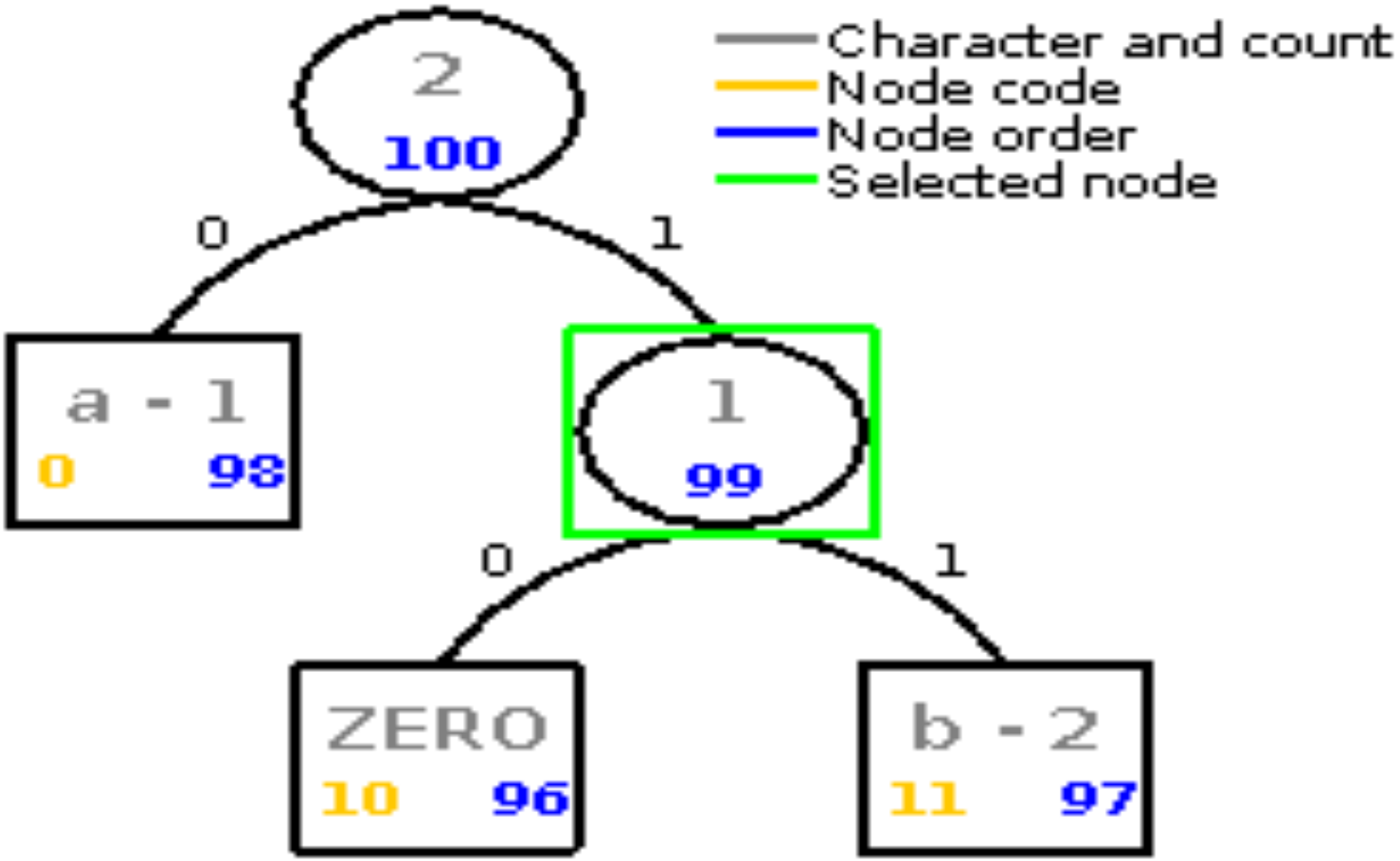
Vitter example: abba



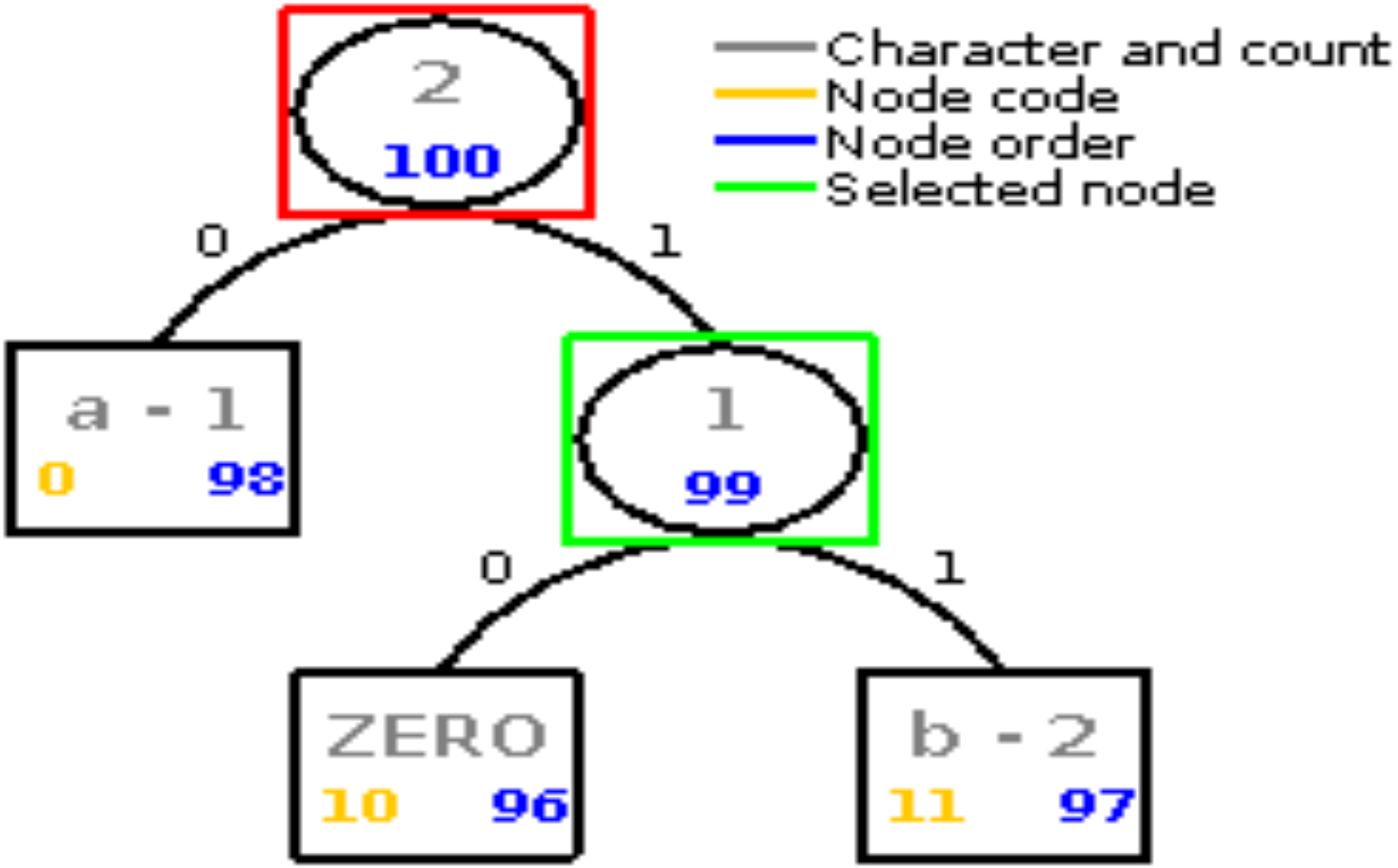
Vitter example: abba



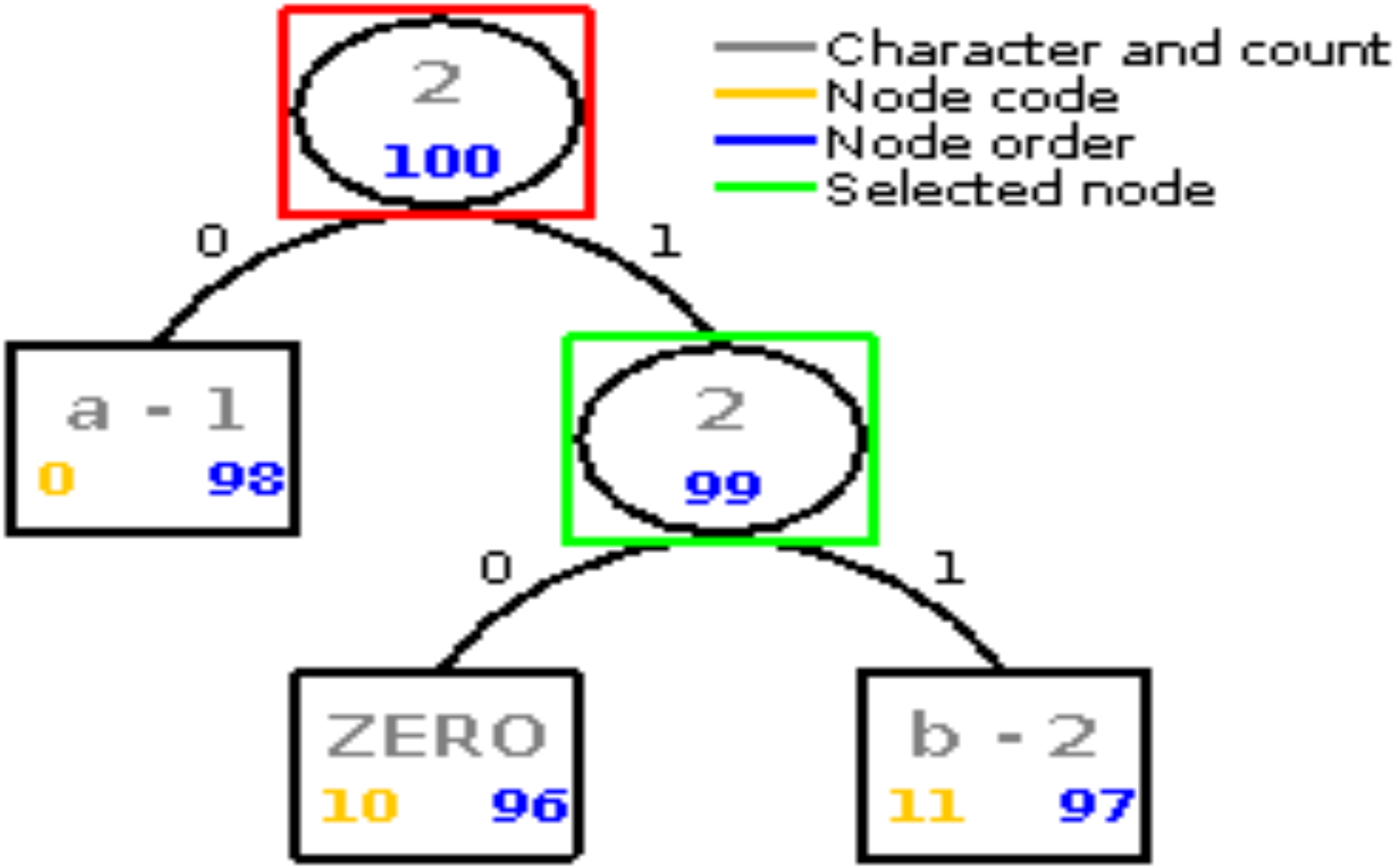
Vitter example: abba



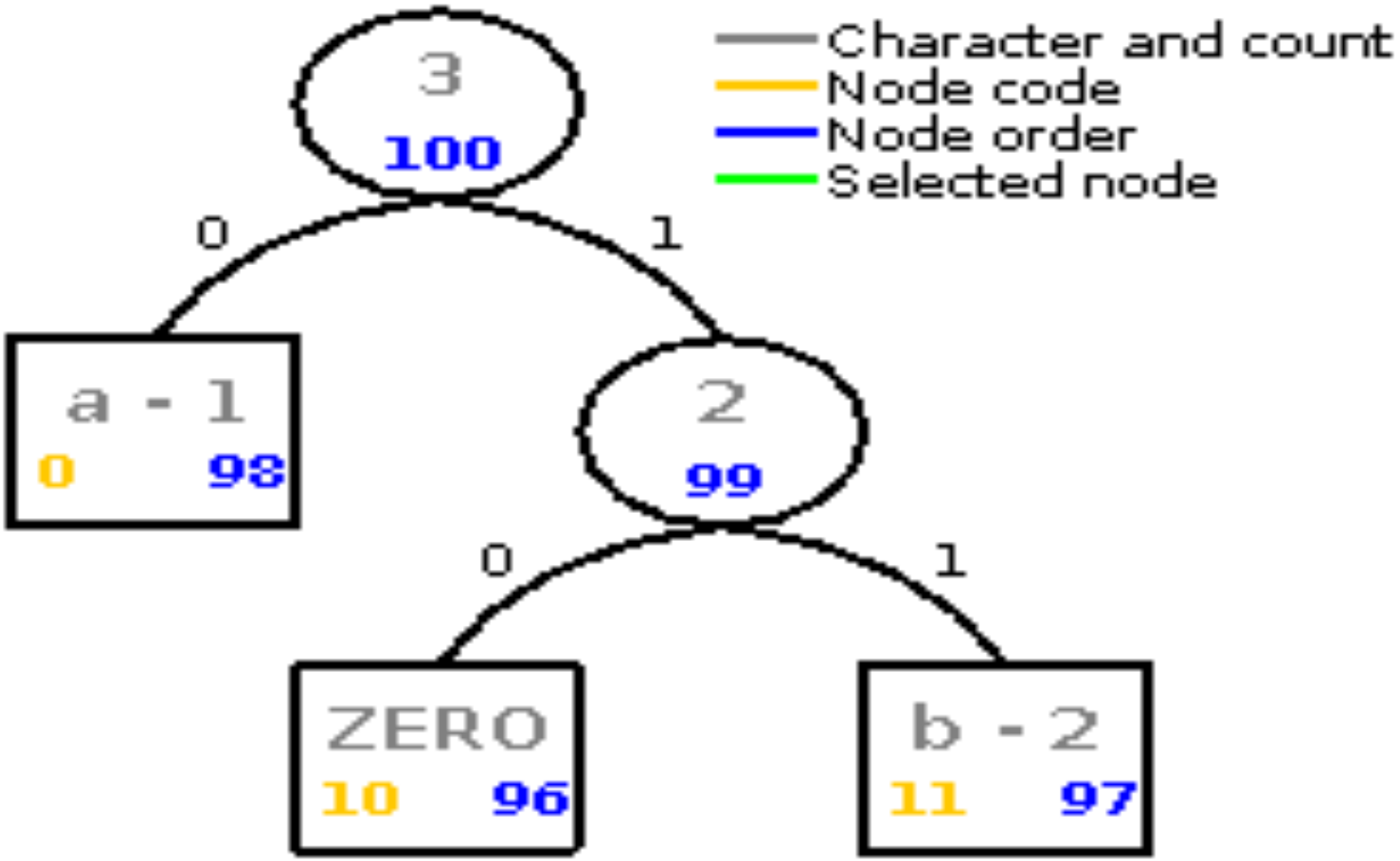
Vitter example: abba



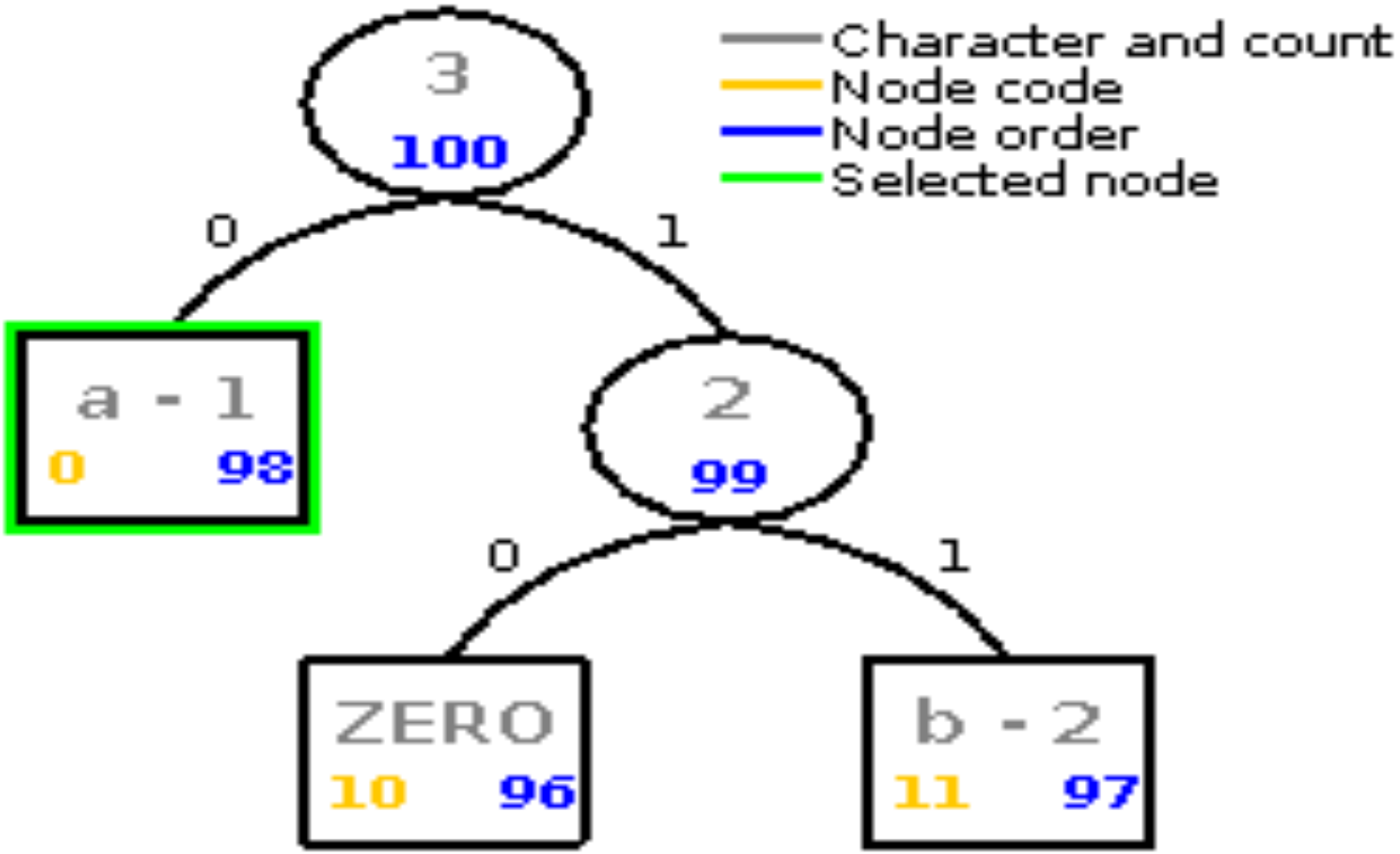
Vitter example: abba



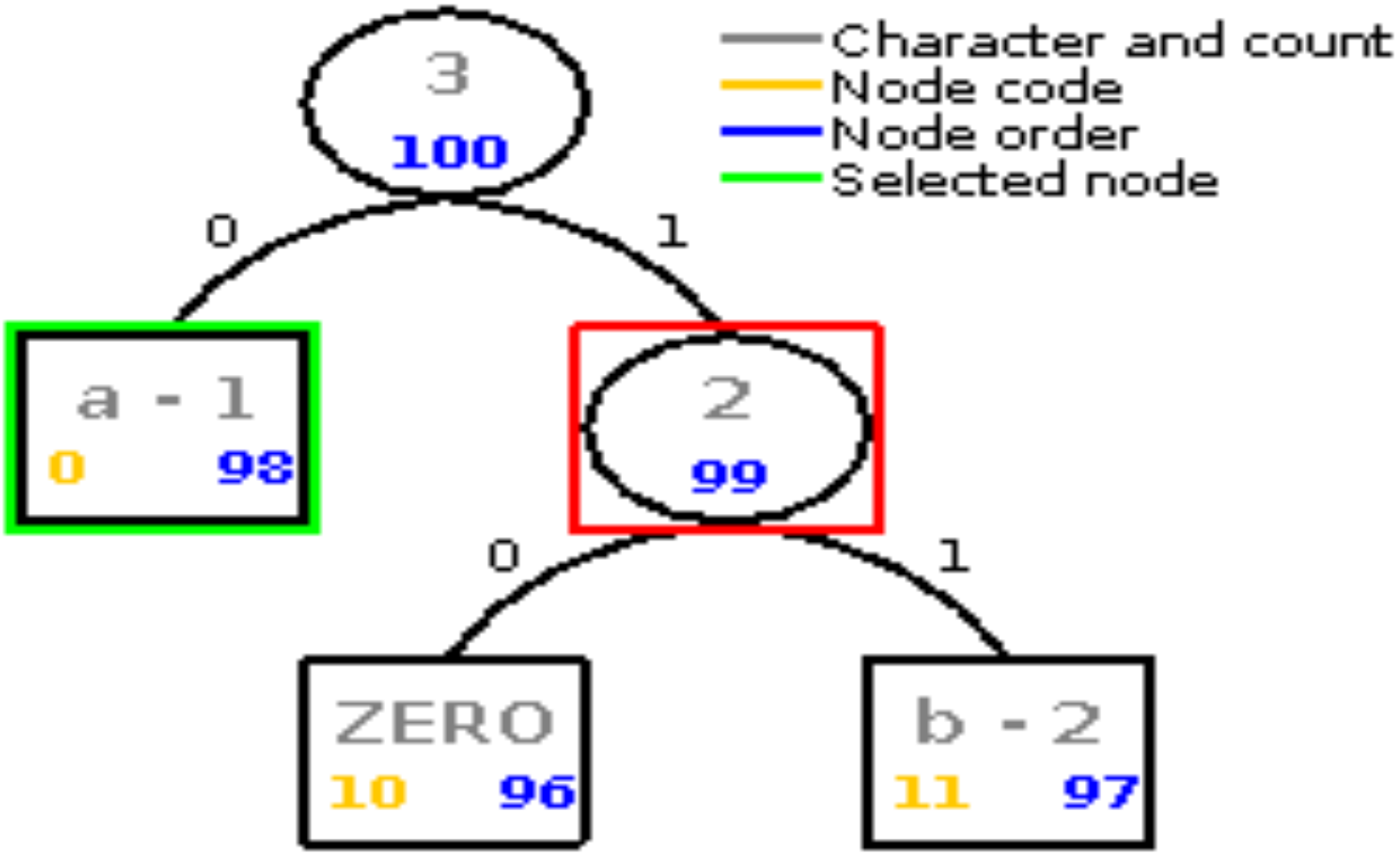
Vitter example: abba



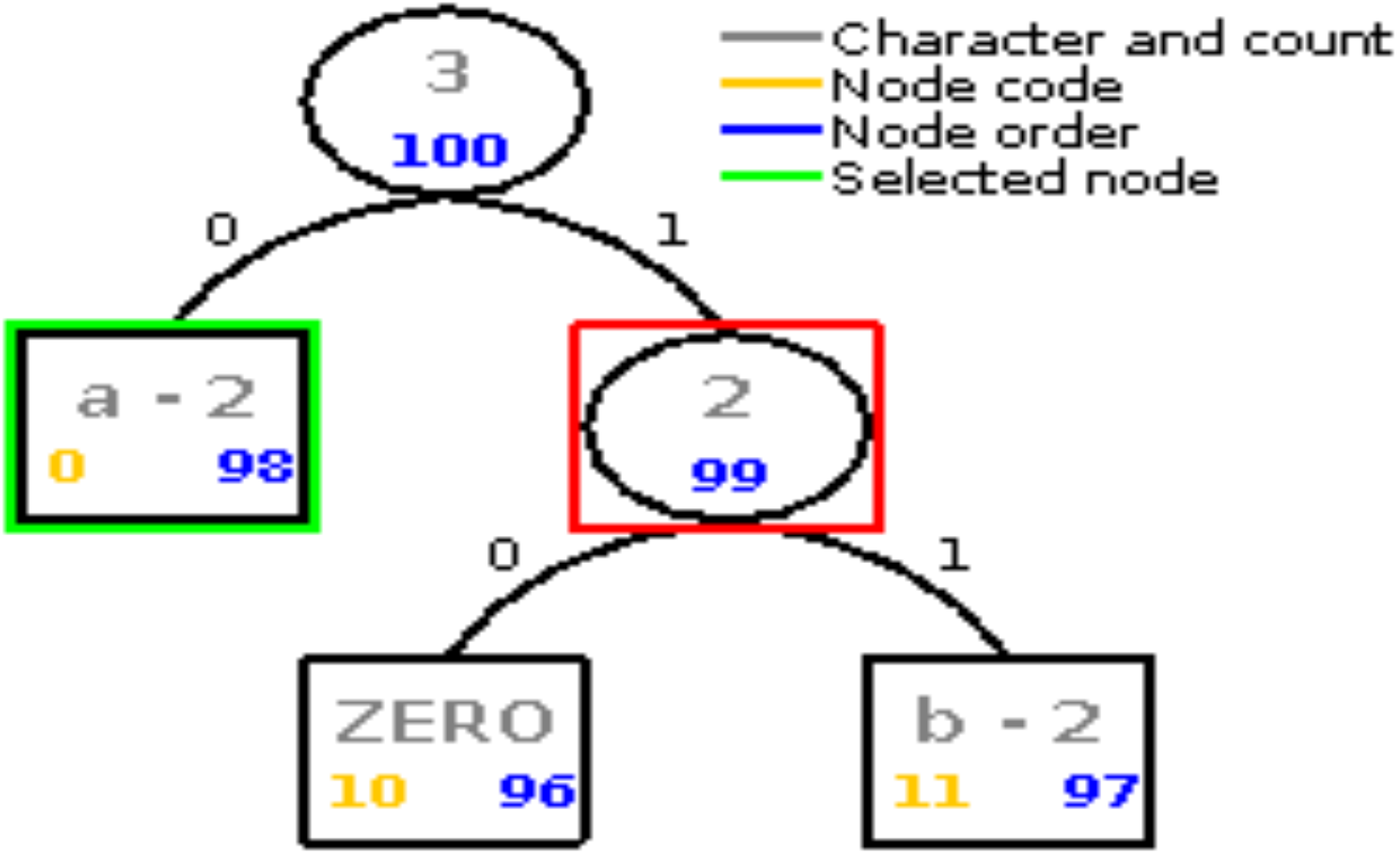
Vitter example: abba



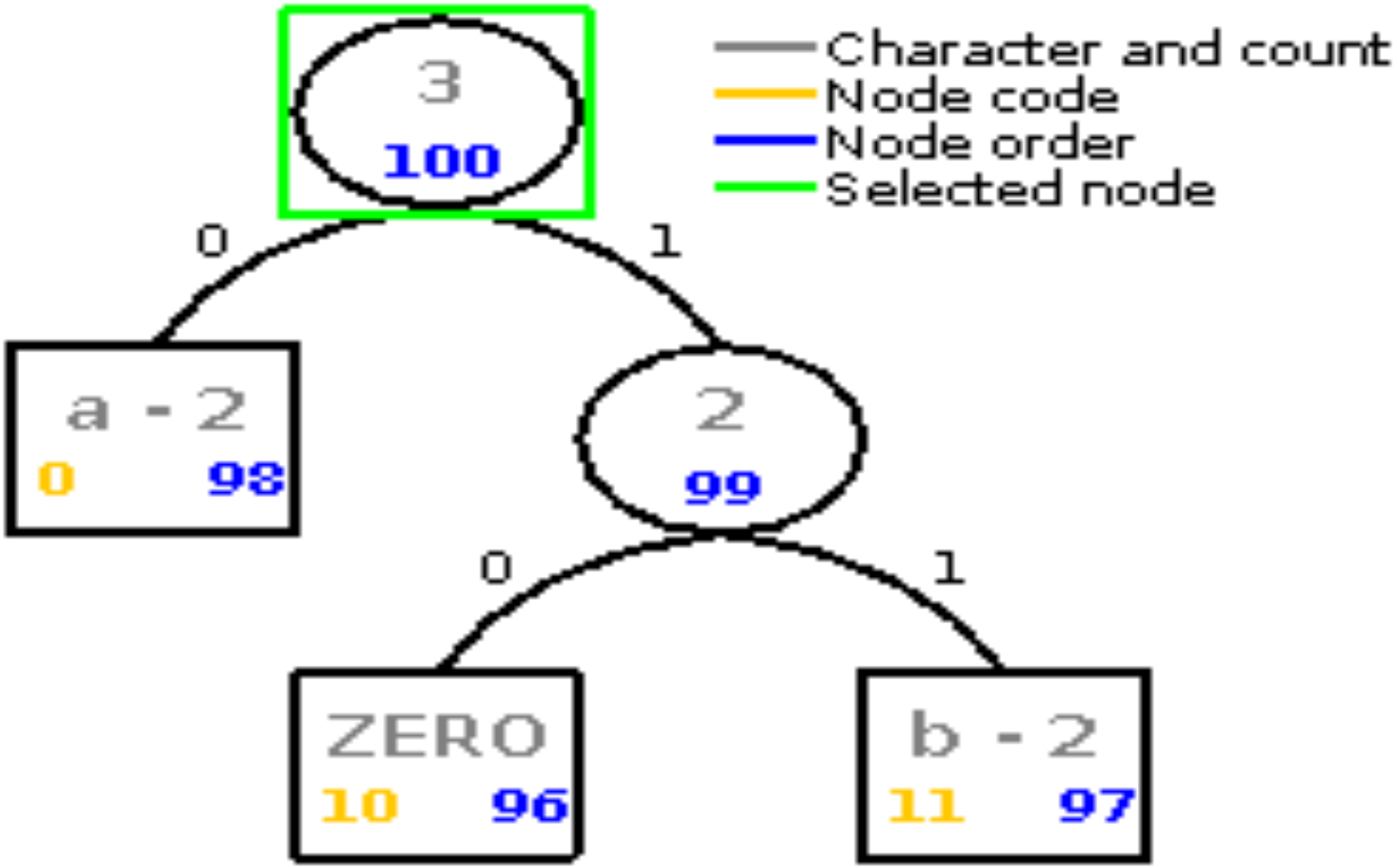
Vitter example: abba



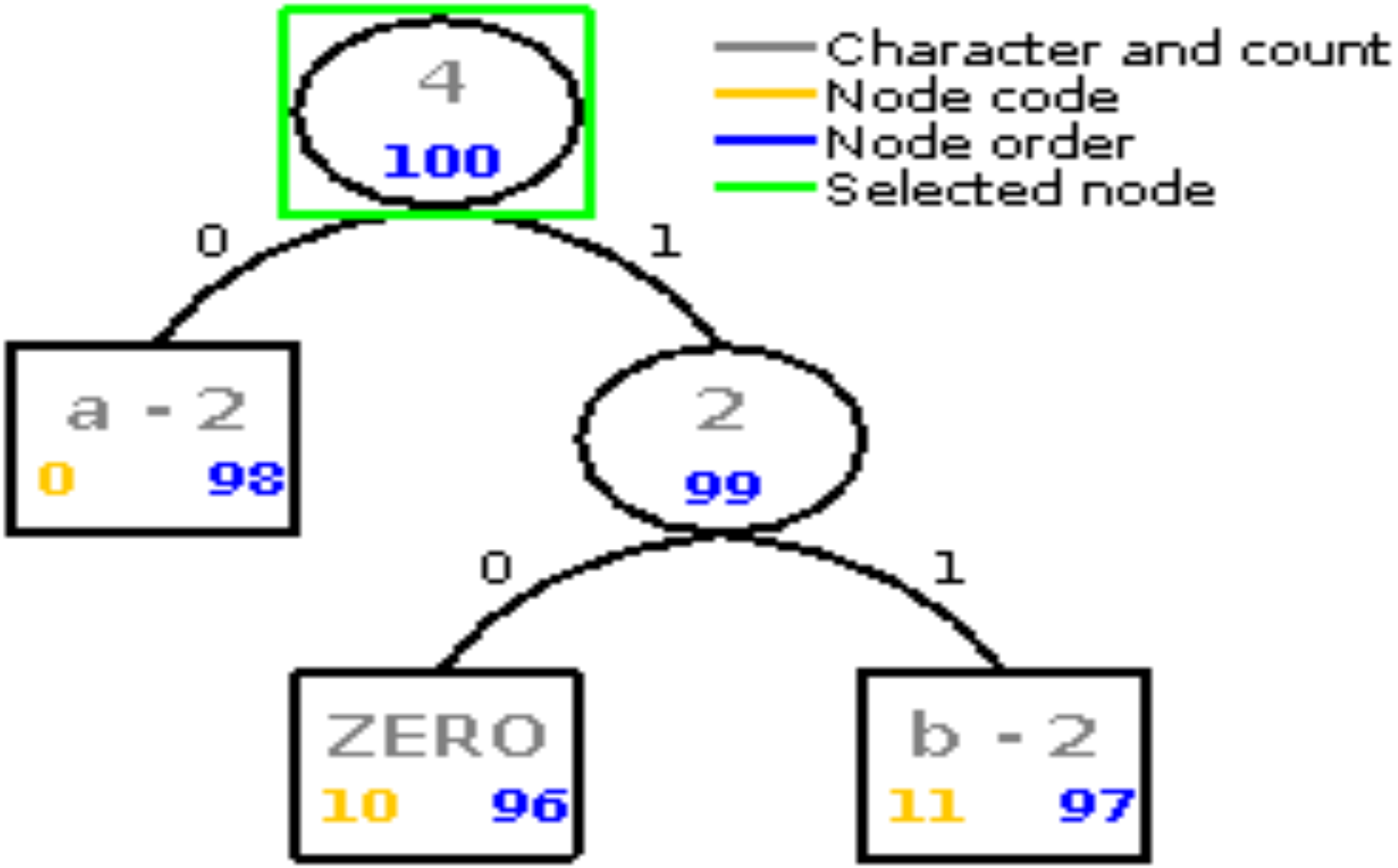
Vitter example: abba



Vitter example: abba



Vitter example: abba



Summary

The main advantage of adaptive Huffman coding is it's ability to code/decode the source string even if is incomplete (e.g. waiting for user input, slow connection), and this can be very helpful for encoding/decoding speed. But one must be aware of it's sensitivity to errors.

References

- readings:
  - http://www.stringology.org/DataCompression/ahv/index_en.html
  - http://www.stringology.org/DataCompression/fgk/index_en.html
  - http://www.cs.duke.edu/csed/curious/compression/adaptivehuff.html#tree
  - http://www.ics.uci.edu/~dan/pubs/DC-Sec4.html
  - http://en.wikipedia.org/wiki/Huffman_coding
- implementation:
  - https://bitbucket.org/matyama/adaptivehuffman