

Bergamini Federico 845646

Errori controllati:

1. Controllo che il numero di campi sia nove per ogni record.
2. Controllo che il campo source sia univoco all'interno del file.
3. Controllo che il campo feature assuma valori compresi in ['5UTR', '3UTR', 'start_codon', 'stop_codon', 'CDS', 'inter', 'inter_cns', 'intron_CNS', 'exon'].
4. Controllo che nel campo attributes siano presenti gene_id e transcript_id.
5. Controllo che il campo seqname sia univoco per ogni gene contenuto nel file.
6. Controllo che i campi start ed end siano contemporaneamente:
 - a. Interi positivi
 - b. $end \geq start$
7. Controllo la validità del campo score, in particolare che sia compreso tra una delle seguenti opzioni:
 - a. Un intero
 - b. Un float
 - c. Un punto
8. Controllo che il campo strand assuma solo valori compresi in ('+', '-').
9. Controllo che il campo strand assuma lo stesso valore per ogni gene contenuto nel file.
10. Controllo che il campo frame assuma solo valori compresi in ('0', '1', '2', '.')
11. Controllo che il campo frame assuma solo valori ammissibili in funzione della feature corrispondente a quel record.
12. Controllo che per ogni trascritto di ogni gene siano presenti le feature start_codon, stop_codon, CDS
13. Per ogni trascritto di ogni gene controllo che gli indici di inizio e di fine degli esoni non si sovrappongano.
14. Per ogni trascritto di ogni gene controllo che gli indici di inizio e di fine dei segmenti di CDS non si sovrappongano.
15. Controllo che le feature start_codon e stop_codon siano della giusta lunghezza, supponendo che potrebbero anche non essere atomiche

N.B. se si verifica un errore, l'esecuzione del validatore si ferma.