PostingsHighlighter

PostingsHighlighter is a new highlighter in ⚠️ Solr4.3 to summarize documents for summary results.

- Introduction
- solrconfig
- schema
- configuration parameters

Introduction

There are already two highlighters, why another?

1. What postingshighlighter is:
    - Uses significantly less disk space than term vectors (~ 1.1byte/position for wikipedia).
    - Passage ranking algorithm focuses on good document summaries.
    - A performant approach if queries have a relatively low number of terms compared to number of results displayed per page.
2. What postingshighlighter is not:
    - Not a query/matching debugger: It just tries to summarize the document with respect to the query terms. If you want to "highlight wildcards", you won't be very happy with this. On the other hand if you want fast highlighting for full-text search, read on.
    - Not for broken analysis chains: When you use `storeOffsetsWithPositions`, IndexWriter enforces that the offsets are correct and won't allow bogus data into the index. This allows for efficient highlighting algorithms and data compression.
    - Not for the risk-adverse: The code is very new and probably still has some exciting bugs!

solrconfig

This is a configuration with all the defaults. All of the configuration can be specified at query-time too, and per-field (e.g. f.text.hl.tag.post=xxxx):

```
<searchComponent class="solr.HighlightComponent" name="highlight">
  <highlighting class="org.apache.solr.highlight.PostingsSolrHighlighter"/>
</searchComponent>

<requestHandler name="standard" class="solr.StandardRequestHandler">
  <lst name="defaults">
    <int name="hl.snippets">1</int>
    <str name="hl.tag.pre">&lt;em&gt;</str>
    <str name="hl.tag.post">&lt;/em&gt;</str>
    <str name="hl.tag.ellipsis">... </str>
    <bool name="hl.defaultSummary">true</bool>
    <str name="hl.encoder">simple</str>
    <float name="hl.score.k1">1.2</float>
    <float name="hl.score.b">0.75</float>
    <float name="hl.score.pivot">87</float>
    <str name="hl.bs.language"></str>
    <str name="hl.bs.country"></str>
    <str name="hl.bs.variant"></str>
    <str name="hl.bs.type">SENTENCE</str>
    <int name="hl.maxAnalyzedChars">10000</int>
  </lst>
</requestHandler>
```

schema

To use this highlighter, you need to store offsets in parallel with the position data in the index.

```
<field name="text" type="text" indexed="true" stored="true" storeOffsetsWithPositions="true"/>
```

configuration parameters

See the javadoc for a full description.