Using Alignment in PBMT

Phrase extraction based on word alignments is wrong:

- From statistical point of view:
  - No link to the decoding, i.e. the use of the phrases in MT.
  - Wuebker et al. (2010) run "forced" or "constraint" decoding on the training data to obtain phrasal alignments.
  - The overfitting to long phrases is avoided by "leaving-one-out" (Ney et al., 1995).
- From linguistic point of view:
  - Fraser and Marcu (2007) allow for M-to-N non-consecutive translation units.
  - DeNero and Klein (2010) train on manual word alignments and handle "possible" links specifically.