

Bölüm 5: Dizgi Algoritmaları

Algoritmalar





- Sonek Dizisi (Suffix Array)
 - Bir dizginin tüm son eklerinin bir dizisi.
 - Dizgi içindeki alt dizgilerin bir temsili olarak kullanılır.
- Burrows-Wheeler Dönüşümü (BWT)
 - Bir dizginin tersine dönüştürülmesiyle elde edilen yeni bir form,
 - Bzip2 gibi sıkıştırma algoritmaları için önişlem adımı olarak kullanılır.





- Bir dizginin tüm soneklerinin alfabetik olarak sıralanmış halidir.
- Her bir sonek, dizginin belirli bir konumundan başlayan bir alt dizisidir.
- Dizgi arama, sıralama, genetik dizilim analizi benzeri işlemlerde kullanılır.





- Brute Force:
 - Tüm sonekleri oluşturup ardından sıralar.
 - O(n² log n) zaman karmaşıklığına sahiptir.
- Manber-Myers Algoritması:
 - Lineer zaman karmaşıklığına sahiptir. (O(n log n))
 - Soneklerin sıralanması sırasında tekrar tekrar harf karşılaştırması yapılmaz.
- Larsson-Sadakane Algoritması:
 - Sıralama işleminde art arda soneklerin karşılaştırılması esas alınır.







6

- Girdi dizisini yeniden düzenler.
- Aynı karakterlerin bir araya toplanmasını sağlar.
- Sıkıştırma algoritmalarının performansını artırır.
- Tersine çevrilebilir: Orijinal veri, dönüşümden geri elde edilebilir.





- Girdi dizgisinin tüm döndürülmüş hallerini oluştur.
- Döndürülmüş dizgileri alfabetik olarak sırala.
- Sıralanmış dizgilerin son karakterlerinden yeni bir dizgi oluştur.

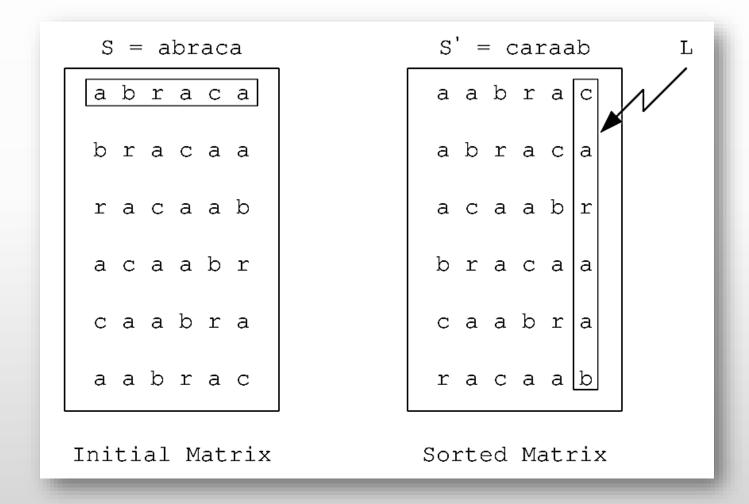
Örnek



- Girdi:
 - BANANA
- Döndürmeler:
 - BANANA, ANANAB, NANABA, ANABAN, NABANA, ABANAN
- Sıralanmış Döndürmeler:
 - ABANAN, ANABAN, ANANAB, BANANA, NABANA, NANABA
- Son Karakterler:
 - NNBAAA
- BWT Sonucu:
 - NNBAAA

Burrows-Wheeler







SON