

درختِ تِرای

trie

# مقدمات

- فرض کنیم می‌خواهیم تمام کلمات انگلیسی را ذخیره کنیم.

- تعداد کلمات انگلیسی حدوداً ۴۷۰,۰۰۰ است

- اگر از یک لیست برای نگهداری کلمات استفاده شود

- هزینه جست و جو چقدر در بدترین حالت چقدر است؟

# مقدمات

- فرض کنیم می‌خواهیم تمام کلمات انگلیسی را ذخیره کنیم.

- تعداد کلمات انگلیسی حدوداً ۴۷۰,۰۰۰ است

- اگر از یک لیست برای نگهداری کلمات استفاده شود

- هزینه جست و جو می‌تواند ۴۷۰,۰۰۰ واحد نیز شود

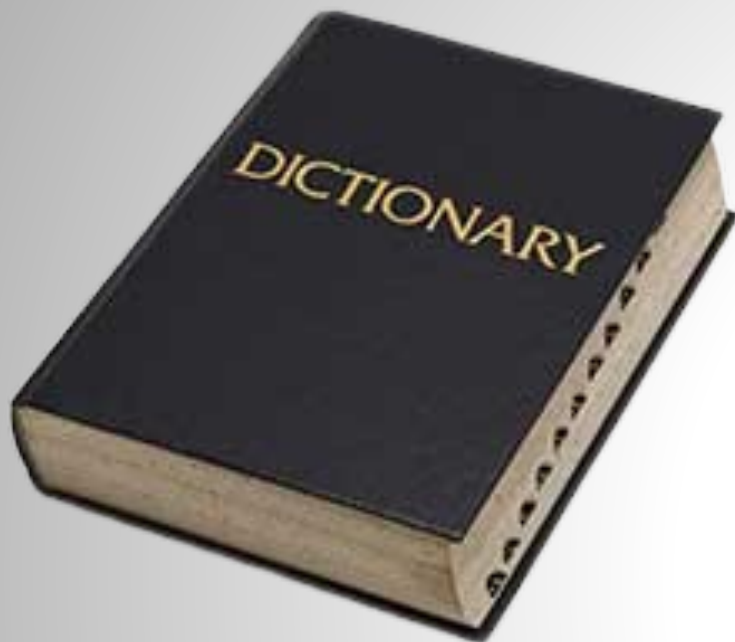
# مقدمات

- فرض کنیم می‌خواهیم تمام کلمات انگلیسی را ذخیره کنیم.

- بزرگترین کلمه انگلیسی ۴۵ حرف دارد

- میانگین طول کلمات انگلیسی ۴.۷ است

# مقدمات



- فرض کنیم می‌خواهیم تمام کلمات انگلیسی را ذخیره کنیم.

- بزرگترین کلمه انگلیسی ۴۵ حرف دارد

- میانگین طول کلمات انگلیسی ۴.۷ است

- می‌خواهیم با ایده‌ای مشابه به کتاب لغتنامه فیزیکی، داده ساختاری با پیچیدگی برحسب طول لغت بسازیم.

- بیشترین هزینه جست و جو: ۴۵

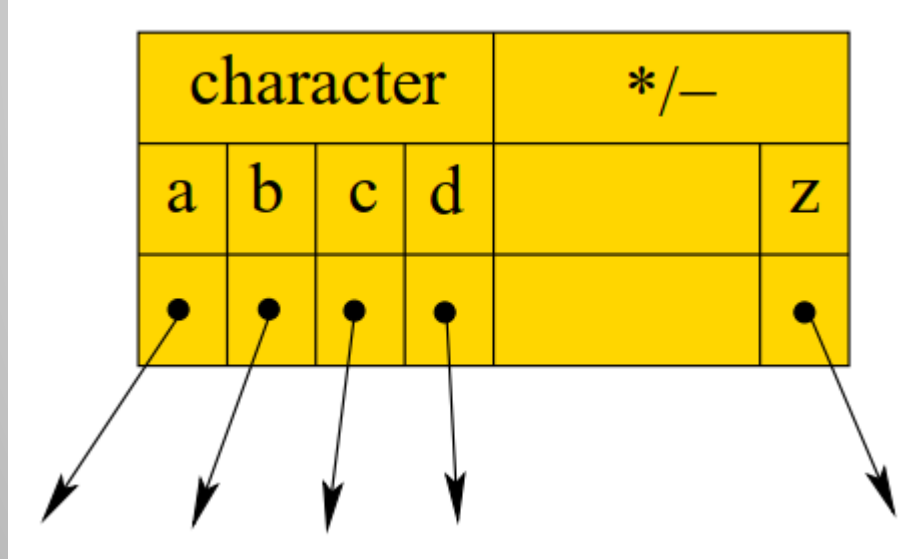
- میانگین (average not expected) هزینه جست و جو: ۴.۷

# ساختمان داده trie

- نام trie برگرفته از کلمه retrieval است اما برای آن که با کلمه tree اشتباه گرفته نشود، آن را try تلفظ می‌کنند.
- این داده ساختار، نوعی درخت است، که هدف آن ذخیره کردن تعداد زیادی داده با طول محدود می‌باشد.

# ساختمان داده trie

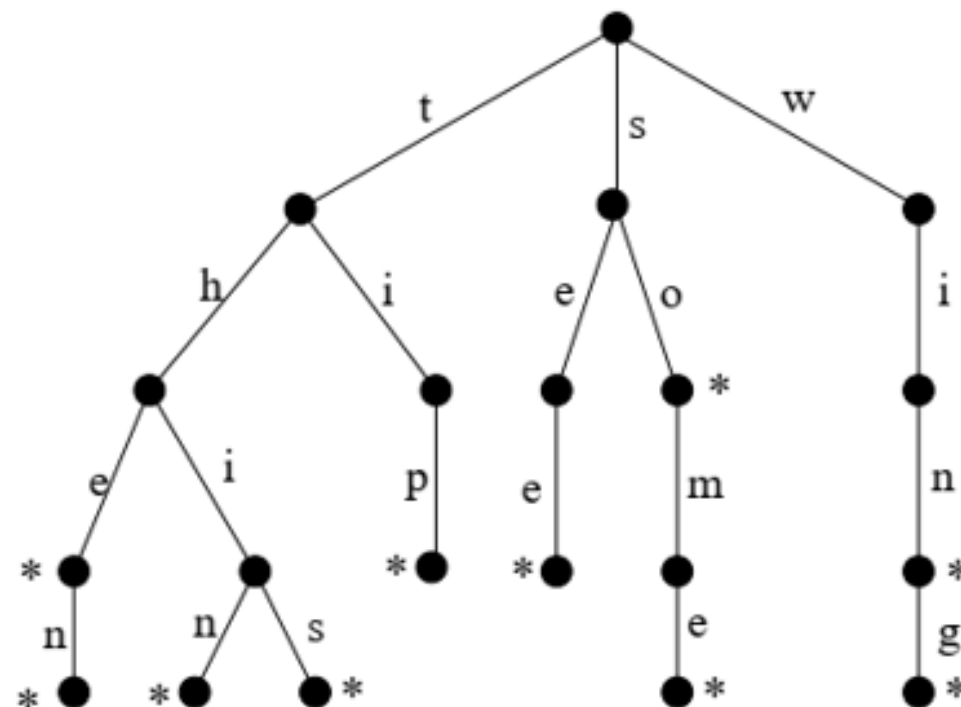
- از آنجایی که زبان انگلیسی ۲۶ حرف دارد، ساختمان داده trie برای آن، یک درخت ۲۶ تایی می‌باشد که هر زیر درخت کلماتی که با حرف یکسان شروع می‌شود را ذخیره می‌کند



- هر گره شامل:
- یک متغیر برای نگهداشتن مقدار مورد نظر (flag)
- ۲۶ اشاره‌گر برای نگهداشتن فرزندان خود

# ساختمان داده trie

• مثال: درخت تری برای عناصر

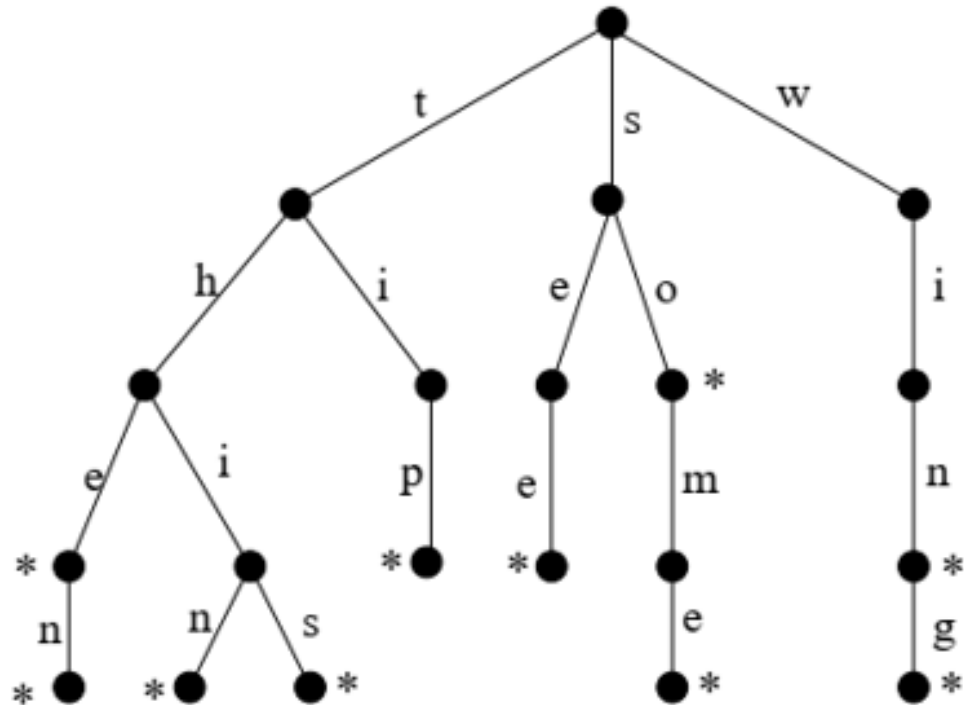


The  
Then  
Thin  
This  
Tip  
See  
So  
Some  
Win  
wing



# ساختمان داده trie

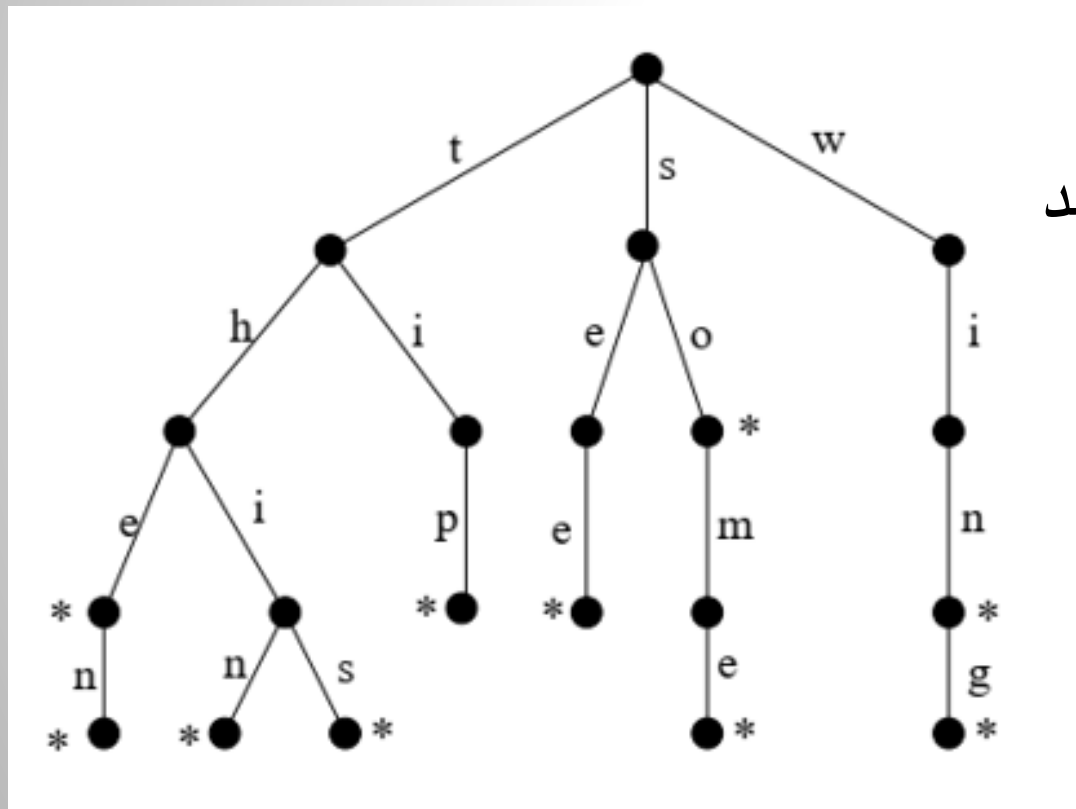
- مقدار هر عنصر برابر با مقدار مشخص کننده مسیر از ریشه تا برگ آن



- اگر دو عنصر prefix یکسانی داشته باشند ویژگی مشترک گره‌های آن چیست؟

# ساختمان داده trie

- مقدار هر عنصر برابر با مقدار مشخص کننده مسیر از ریشه تا برگ آن



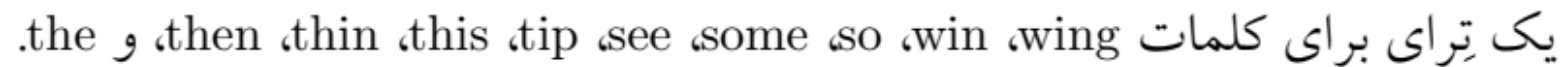
- اگر دو عنصر prefix یکسانی داشته باشند  
جد آنها یکسان است

# اعمال مورد نیاز روی داده ساختار trie

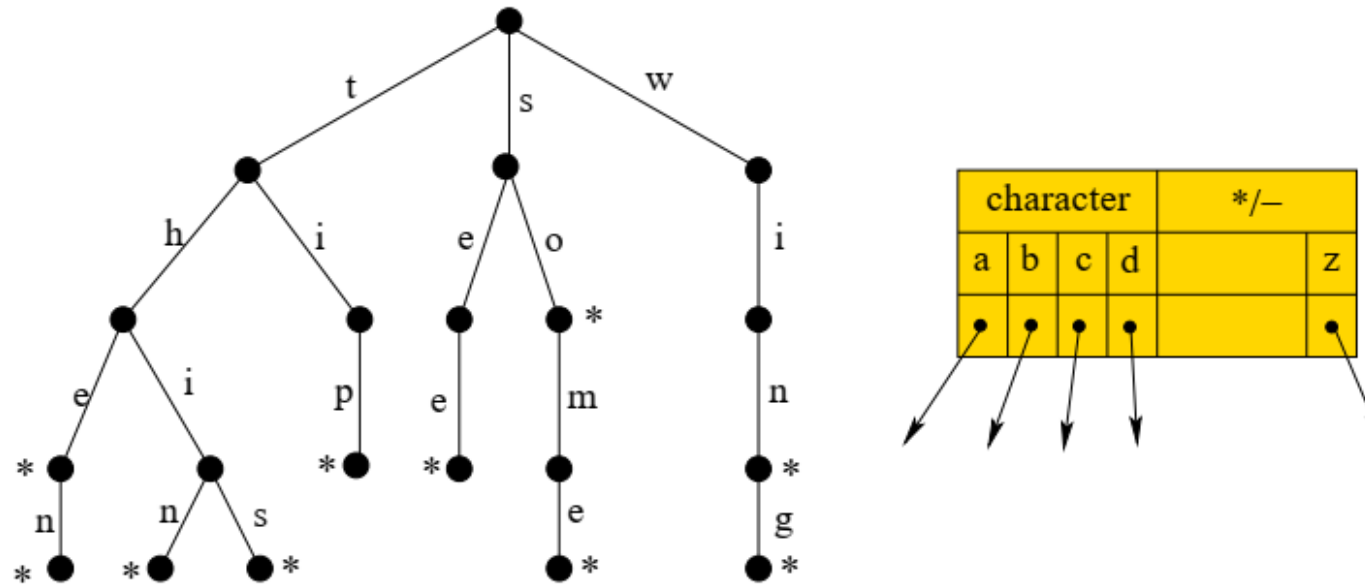
- اضافه کردن یک رشته

- جست و جو یک رشته

- حذف کردن یک رشته

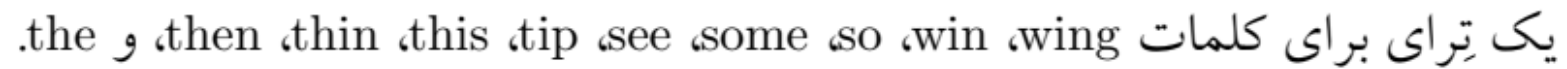


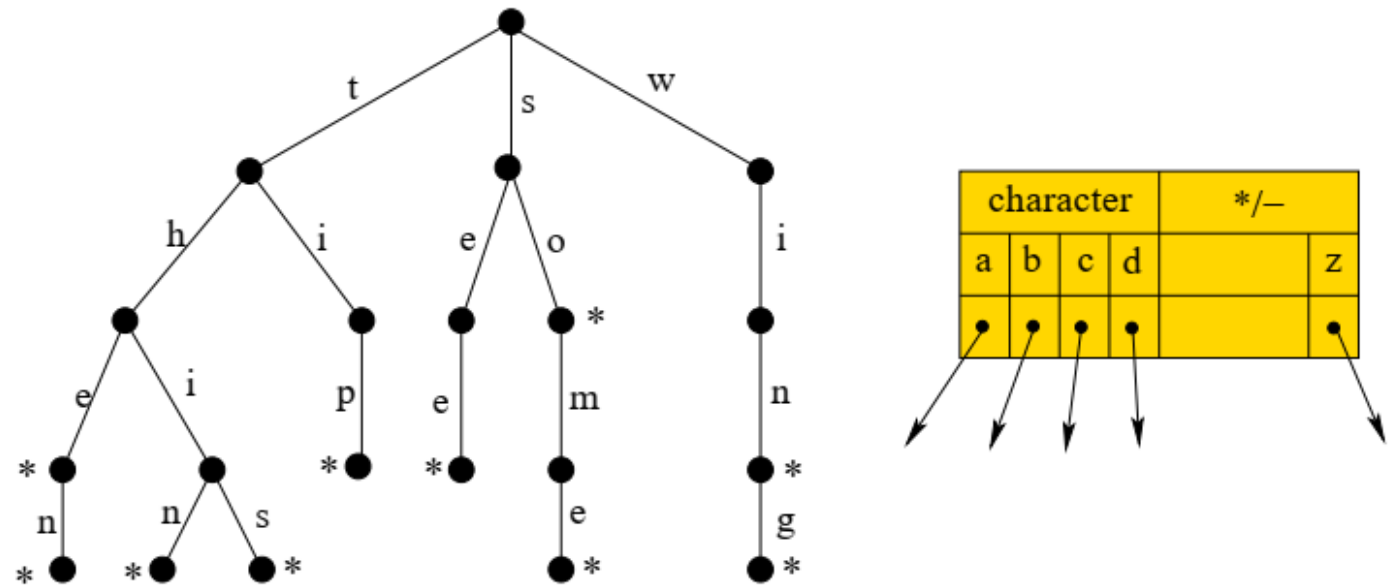
## اضافه کردن will



یک تِرای برای کلمات the, then, thin, this, tip, see, some, so, win, wing و .the

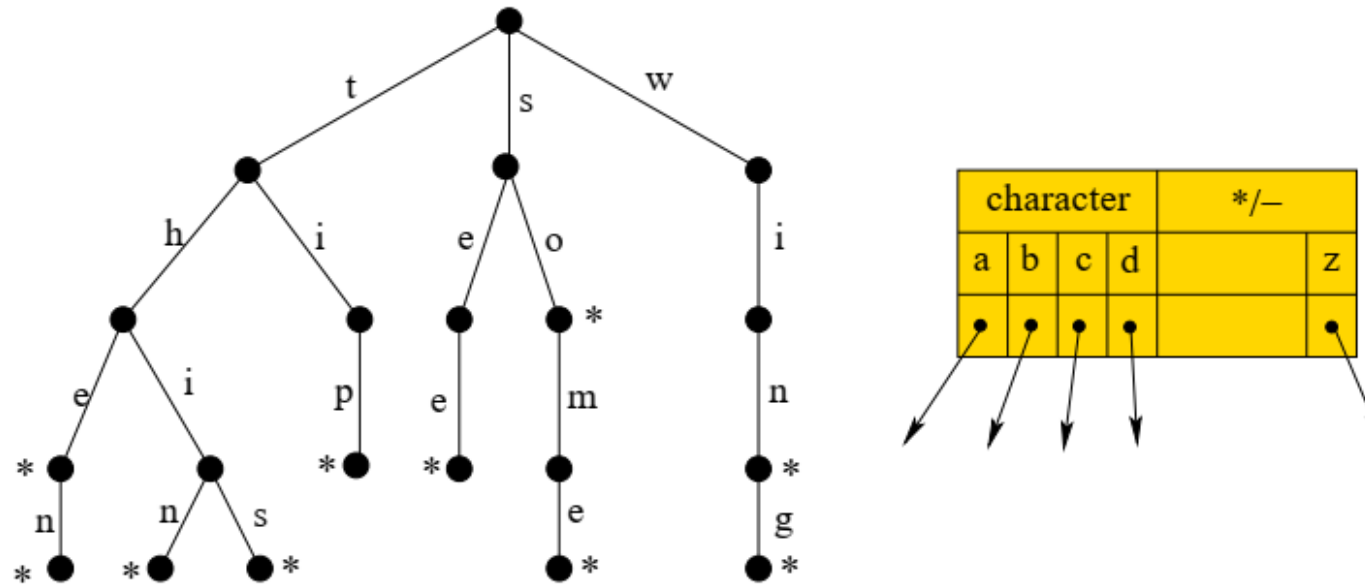




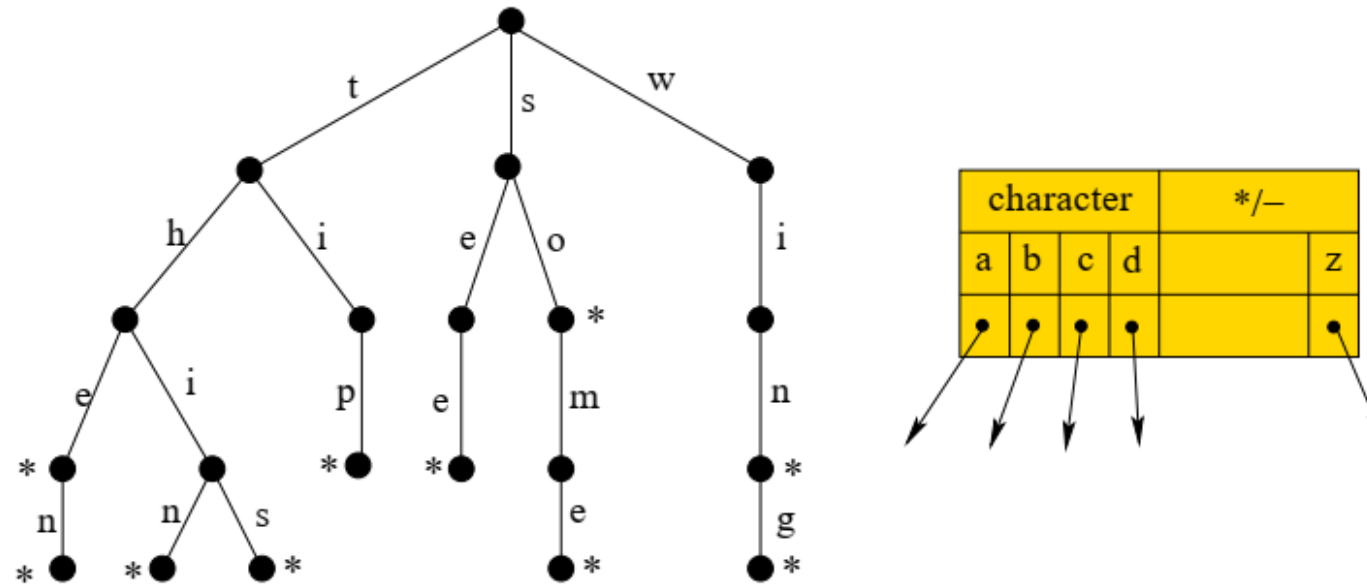


یک تِرای برای کلمات the, then, thin, this, tip, see, some, so, win, wing





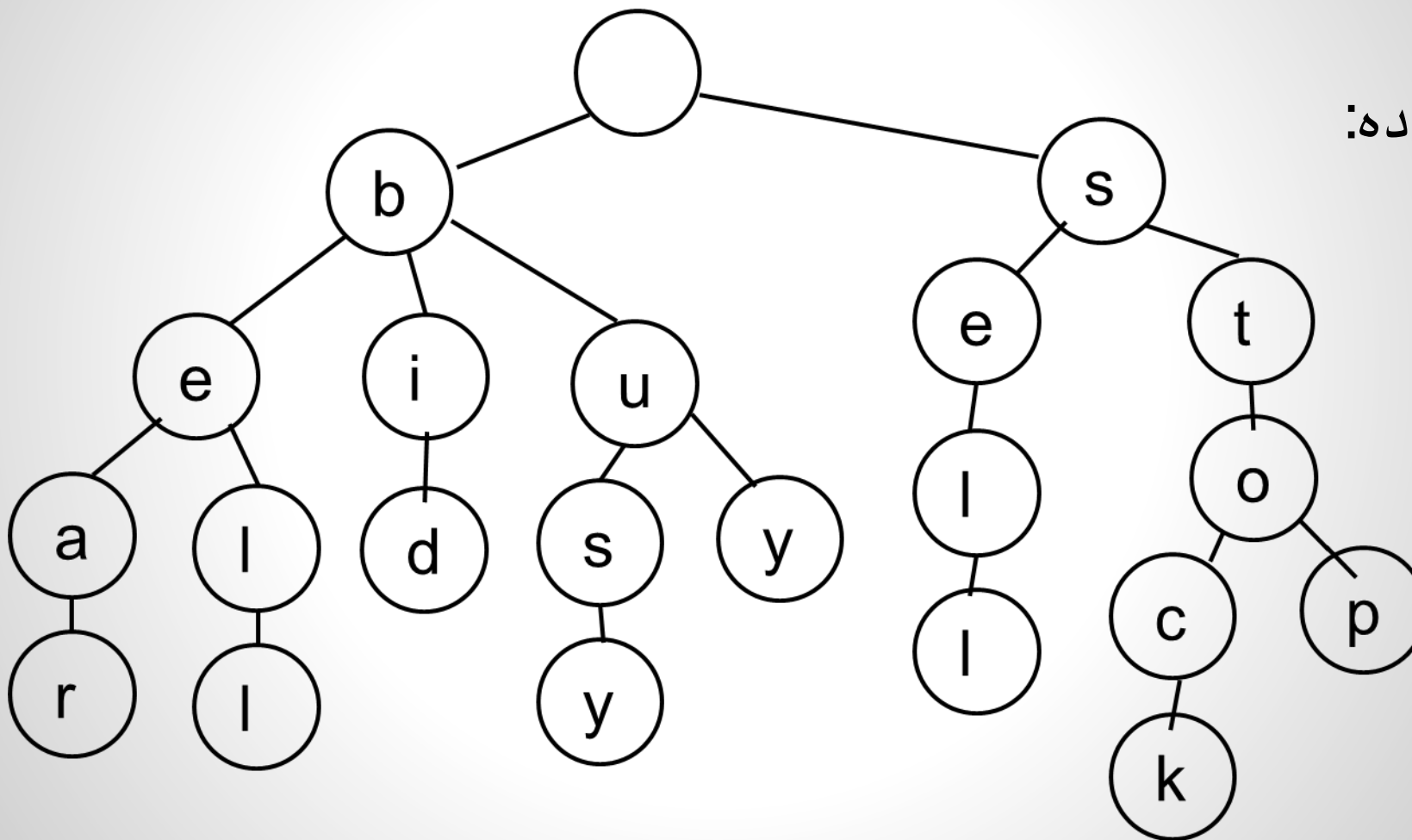
یک تِرای برای کلمات the, then, thin, this, tip, see, some, so, win, wing



یک تِرای برای کلمات the, then, thin, this, tip, see, some, so, win, wing و .the

# فشرده سازی (در امتحان نمیاد)

• درخت ساده:



# فشرده سازی (در امتحان نمیاد)

• درخت فشرده با ادغام گره‌های تک فرزند با فرزندشان

