

Parts of strings

1 Substrings

A **substring** is a continuous part of a string.

EXAMPLE 1.

The string *abcd* has 11 substrings:

- ε
- *a*
- *b*
- *c*
- *d*
- *ab*
- *bc*
- *cd*
- *abc*
- *bcd*
- *abcd*

Note that

1. the empty string is a substring of every string, and
2. every string is a substring of itself.

A substring *u* of *v* is **proper** iff $u \neq v$.

EXAMPLE 2.

All the strings listed above are proper substrings of *abcd*, except *abcd* itself.

EXERCISE 1.

For each one of the gaps below, enter \sqsubseteq , \sqsubset , or $\not\sqsubseteq$ depending on whether the first string is a substring of the second string, a proper substring, or neither:

- *a_aaaa*
- *a_b*
- ε_b
- ε_ε
- *aa_abbbca*
- *bc_abbbca*
- *cb_abbbca*

2 Subsequence

A **subsequence** is a discontinuous part of a string that preserves the order between the symbols.

EXAMPLE 3.

The string $abcd$ has subsequences:

- ε
- a
- b
- c
- d
- ab
- ac
- ad
- bc
- bd
- cd
- abc
- abd
- bcd
- $abcd$

Note that ca is not a subsequence of $abcd$, but it is a subsequence of $abcda$.

Just like substrings, a subsequence u of v is proper iff $u \neq v$.

EXERCISE 2.

For each one of the gaps below, enter \sqsubseteq , \subsetneq , or $\not\sqsubseteq$ depending on whether the first string is a subsequence of the second string, a proper subsequence, or neither:

- a_aaaa
- a_b
- ε_b
- $\varepsilon_ \varepsilon$
- aa_abbbca
- bc_abbbca
- cb_abbbca

EXERCISE 3.

Say whether the following is True or False: Every substring of some string s is also

a subsequence of s , but not the other way round. Justify your answer.