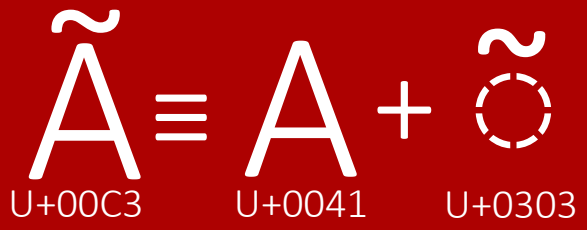
# Unicode String Normalization

10084




Use Cases:

- Internet/Web Protocols: IDNA, PRECIS,…
- User name/password entry, text processing

Proposal: Built-in (or Standard Library)

`'string'.normalize/.nfc` ⇒ NFC version of `'string'`

`'string'.nfkd` ⇒ NFKD version of `'string'`

Pros:

- Natural OO syntax (not e.g. `UnicodeUtils.nfkd 'string'`)
- Libraries available:
  - *Pure Ruby: eprun (https://github.com/duerst/eprun),…*
  - *C: unf,… (http://bibwild.wordpress.com/2013/11/19/benchmarking-ruby-unicode-normalization-alternatives/)*

Cons: Unicode version updates needed; what about non-UTF-8 strings?