Matando (ou quase) Unicode (De | En) code Errors No Python 2

Para acabar (ou quase) com os erros é necessário compreender o funcionamento

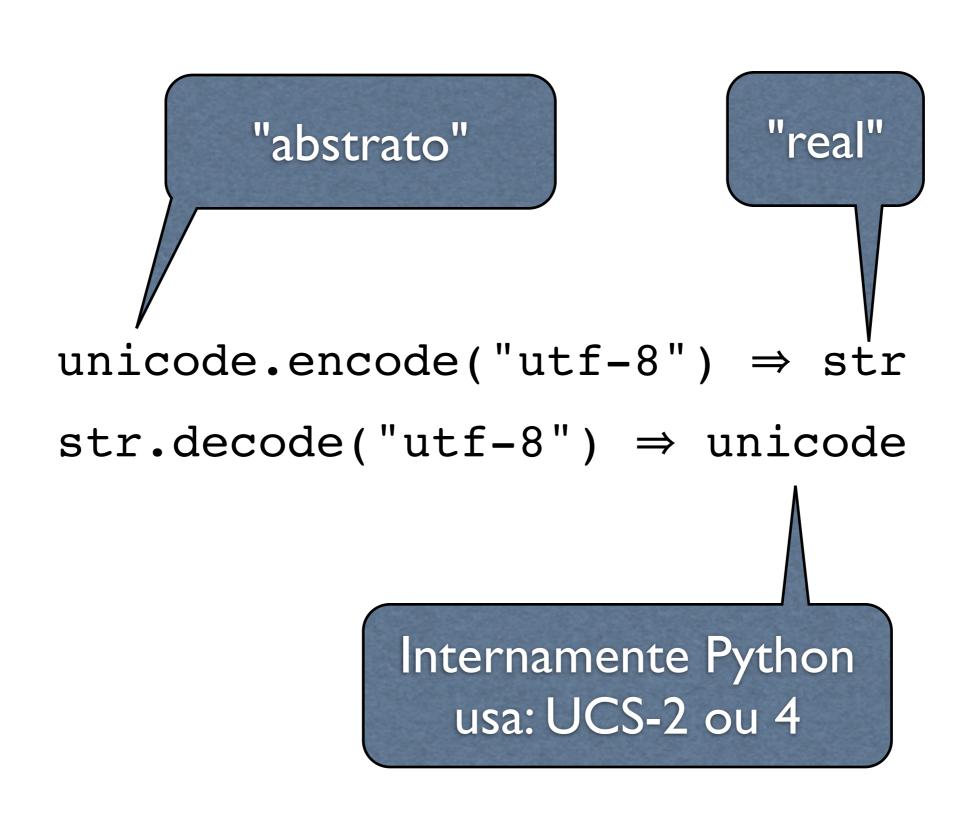
Unicode: catálogo com todos os símbolos criados pela humanidade

Unicode é uma abstração

código numérico ⇒ símbolo

Computadores são reais e não abstratos, então...

Codificação / Encoding (UCS, UTF, ISO, ...)



Unicode é complicado

Codificações são complicadas

Python + Unicode + Codificações = CHAOS

Adivinhar a codificação: impossível

Técnicas para descobrir

- Tentativa e erro
- BOM
- Accept-Charset
- <form accept-charset="???">
- Mais algumas outras...
- Use mais de uma abordagem

Leitura Recomendada

 http://www.python.org.br/wiki/ TudoSobrePythoneUnicode