import unicodedata

# Function to check for mixed scripts

def is\_mixed\_script(url):

scripts = set()

for char in url:

if char.isalpha():

script = unicodedata.name(char).split()[0] # Get the script name

scripts.add(script)

return len(scripts) > 1 # More than one script indicates a potential homograph

# Function to check for confusable characters

def has\_confusable\_char(url):

confusable\_chars = {

'a': ['а'], 'c': ['с'], 'e': ['е'], 'h': ['н'], 'i': ['і'], 'o': ['о'],

'p': ['р'], 's': ['ѕ'], 't': ['т'], 'x': ['х'], 'y': ['у']

}

for char in url:

if char in confusable\_chars:

return True

for similar in confusable\_chars.get(char, []):

if similar in url:

return True

return False

# Main function to check for homographs

def is\_homograph\_url(url):

mixed\_script = is\_mixed\_script(url)

confusable = has\_confusable\_char(url)

return mixed\_script, confusable

# Example usage

url = input("Enter a URL to check for homographs: ")

mixed\_script, confusable = is\_homograph\_url(url)

if mixed\_script and confusable:

print("The URL contains both mixed scripts and confusable characters.")

elif mixed\_script:

print("The URL contains mixed scripts.")

elif confusable:

print("The URL contains confusable characters.")

else:

print("The URL does not contain homograph characters.")



