

The-Compiler/hypothesis-talk
Florian Bruhin
October 17th, 2024







#### About me

The-Compiler/hypothesis-talk









2011

2013, 2015

2019, 2020

#### Calculator project

```
— rpncalc/utils.py
def calc(a, b, op):
    if op == "+":
        return a + b
    elif op == "-":
        return a - b
    elif op == "*":
        return a * b
    elif op == "/":
        return a / b
    raise ValueError("Invalid operator")
```

History



- Using a calculator without needing a = key, and without parentheses
- Makes it much easier to implement, using a stack data structure
- Based on the Polish notation, invented by Jan Łukasiewicz in 1924
- Used by all HP calculators in the 1970s–80s, still used by some today
- Displayed here: HP 12C financial calculator, introduced in 1981, still in production today (HPs longest and best-selling product)

```
Reverse Polish Notation (RPN) Explanation
```

$$1 + 2$$













Reverse Polish Notation (RPN) Explanation

$$5 \cdot (1+2)$$

```
Reverse Polish Notation (RPN)
 — rpncalc/rpn_v1.py -
 from rpncalc.utils import calc
 class RPNCalculator:
     def __init__(self) -> None:
          self.stack = \Pi
     def run(self) -> None:
          while True:
              inp = input("> ")
              if inp == "q":
                  return
              elif inp == "p":
                  print(self.stack)
              else:
                  self.evaluate(inp)
```

```
Reverse Polish Notation (RPN)
 — rpncalc/rpn_v1.py -
 from rpncalc.utils import calc
 class RPNCalculator:
     def __init__(self) -> None:
          self.stack = \Pi
     def run(self) -> None:
          while True:
             inp = input("> ")
             if inp == "q":
                  return
             elif inp == "p":
                  print(self.stack)
              else:
                  self.evaluate(inp)
```

```
Reverse Polish Notation (RPN)
 — rpncalc/rpn_v1.py -
                                            def evaluate(self, inp: str):
 from rpncalc.utils import calc
                                                if inp.isdigit():
                                                    n = float(inp)
 class RPNCalculator:
                                                     self.stack.append(n)
     def init (self) -> None:
                                                 elif inp in "+-*/":
          self.stack = []
                                                    b = self.stack.pop()
                                                    a = self.stack.pop()
     def run(self) -> None:
                                                    res = calc(a, b, inp)
         while True:
                                                     self.stack.append(res)
              inp = input("> ")
                                                     print(res)
              if inp == "q":
                  return
              elif inp == "p":
                                        if __name__ == "__main__":
                  print(self.stack)
                                            rpn = RPNCalculator()
              else:
                                            rpn.run()
                  self.evaluate(inp)
```

```
Reverse Polish Notation (RPN)
 — rpncalc/rpn_v1.py -
                                            def evaluate(self, inp: str):
 from rpncalc.utils import calc
                                                if inp.isdigit():
                                                    n = float(inp)
 class RPNCalculator:
                                                     self.stack.append(n)
     def init (self) -> None:
                                                 elif inp in "+-*/":
          self.stack = []
                                                    b = self.stack.pop()
                                                    a = self.stack.pop()
     def run(self) -> None:
                                                    res = calc(a, b, inp)
         while True:
                                                     self.stack.append(res)
              inp = input("> ")
                                                    print(res)
              if inp == "q":
                  return
              elif inp == "p":
                                        if __name__ == "__main__":
                  print(self.stack)
                                            rpn = RPNCalculator()
              else:
                                            rpn.run()
                  self.evaluate(inp)
```

```
Reverse Polish Notation (RPN)
 — rpncalc/rpn_v1.py -
                                            def evaluate(self, inp: str):
 from rpncalc.utils import calc
                                                 if inp.isdigit():
                                                    n = float(inp)
 class RPNCalculator:
                                                     self.stack.append(n)
     def init (self) -> None:
                                                elif inp in "+-*/":
          self.stack = []
                                                    b = self.stack.pop()
                                                    a = self.stack.pop()
     def run(self) -> None:
                                                    res = calc(a, b, inp)
         while True:
                                                     self.stack.append(res)
              inp = input("> ")
                                                     print(res)
              if inp == "q":
                  return
              elif inp == "p":
                                        if __name__ == "__main__":
                  print(self.stack)
                                            rpn = RPNCalculator()
              else:
                                            rpn.run()
                  self.evaluate(inp)
```

```
Reverse Polish Notation (RPN)
 — rpncalc/rpn_v1.py -
 from rpncalc.utils import calc
 class RPNCalculator:
     def init (self) -> None:
         self.stack = []
     def run(self) -> None:
         while True:
             inp = input("> ")
             if inp == "q":
                 return
             elif inp == "p":
                 print(self.stack)
                                            rpn = RPNCalculator()
             else:
                                            rpn.run()
                  self.evaluate(inp)
```

```
def evaluate(self, inp: str):
        if inp.isdigit():
            n = float(inp)
            self.stack.append(n)
        elif inp in "+-*/":
            b = self.stack.pop()
            a = self.stack.pop()
            res = calc(a, b, inp)
            self.stack.append(res)
            print(res)
if __name__ == "__main__":
```

Towards an improved version

- Fix bugs and add additional error handling:
  - Allow negative numbers and floating-point inputs, not just .isdigit()
  - Fix +- being treated as valid input due to elif inp in "+-\*/":
  - Print error when using an invalid operator
  - Handle ZeroDivisionError when dividing by zero, and IndexError with < 2 elements on stack</li>
- Allow passing a Config object to the calculator, with a custom prompt, instead of "> "
- Support multiple inputs on one line, e.g. 2 1 +

Code: Fixing bugs, improving error handling

```
— rpncalc/rpn _v1 .py ———
                                — rpncalc/rpn _v2 .py -
def evaluate(self, inp: str):
                                def evaluate(self, inp: str) -> None:
    if inp.isdigit():
                                    trv:
        n = float(inp)
                                        self.stack.append(float(inp))
        self.stack.append(n)
                                        return
    elif inp in "+-*/":
                                    except ValueError:
                                        pass
                                    if inp not in ["+", "-", "*", "/"]:
mathspp.com/blog/pydonts/
```

LBYL: Look before you leap

eafp-and-lbyl-coding-styles

**EAFP**: It's easier to ask for forgivenness, than for permission

Code: Fixing bugs, improving error handling

```
def evaluate(self, inp: str):
  if inp.isdigit():
     n = float(inp)
     self.stack.append(n)
  elif inp in "+-*/":
```

```
def evaluate(self, inp: str) -> None:
    try:
        self.stack.append(float(inp))
        return
    except ValueError:
        pass
```

if inp not in ["+", "-", "\*", "/"]:

```
Code: Fixing bugs, improving error handling
— rpncalc/rpn_v2.pv -
def err(self, msg: str) -> None:
                                            b = self.stack.pop()
     print(msg, file=sys.stderr)
                                            a = self.stack.pop()
def evaluate(self, inp: str) -> None:
                                            try:
                                                res = calc(a, b, inp)
     if inp not in ["+", "-", "*", "/"]:
                                            except ZeroDivisionError:
         self.err(
                                                self.err("Division by zero")
             f"Invalid input: {inp}")
                                                return
         return
                                            self.stack.append(res)
     if len(self.stack) < 2:
                                            print(res)
         self.err("Not enough operands")
         return
```

```
Code: Fixing bugs, improving error handling
— rpncalc/rpn_v2.pv -
def err(self, msg: str) -> None:
                                            b = self.stack.pop()
     print(msg, file=sys.stderr)
                                            a = self.stack.pop()
def evaluate(self, inp: str) -> None:
                                            try:
                                                res = calc(a, b, inp)
     if inp not in ["+", "-", "*", "/"]:
                                            except ZeroDivisionError:
         self.err(
                                                self.err("Division by zero")
             f"Invalid input: {inp}")
                                                return
         return
                                            self.stack.append(res)
     if len(self.stack) < 2:
                                            print(res)
         self.err("Not enough operands")
         return
```

```
def evaluate(self, inp: str) -> None:
    if inp not in ["+", "-", "*", "/"]:
        self.err(
            f"Invalid input: {inp}")
        return
```

```
if len(self.stack) < 2:
    self.err("Not enough operands")
    return</pre>
```

```
b = self.stack.pop()
a = self.stack.pop()
```

```
res = calc(a, b, inp)
except ZeroDivisionError:
    self.err("Division by zero")
    return
```

self.stack.append(res)

try:

print(res)

```
Code: Fixing bugs, improving error handling
— rpncalc/rpn_v2.pv
def err(self, msg: str) -> None:
     print(msg, file=sys.stderr)
def evaluate(self, inp: str) -> None:
     if inp not in ["+", "-", "*", "/"]:
         self.err(
             f"Invalid input: {inp}")
         return
                                            self.stack.append(res)
     if len(self.stack) < 2:
                                            print(res)
         self.err("Not enough operands")
         return
```

```
b = self.stack.pop()
a = self.stack.pop()
trv:
   res = calc(a, b, inp)
except ZeroDivisionError:
    self.err("Division by zero")
    return
```



## Property-based testing with Hypothesis Motivation

From the Hypothesis website, hypothesis.works:

Most testing is ineffective

Normal "automated" software testing is surprisingly manual. Every scenario the computer runs, someone had to write by hand. Hypothesis can fix this.

# Property-based testing with Hypothesis Motivation

From the Hypothesis website, hypothesis.works:

#### Most testing is ineffective

Normal "automated" software testing is surprisingly manual. Every scenario the computer runs, someone had to write by hand. Hypothesis can fix this.

Idea behind "Property-Based testing" (popularized by QuickCheck/Haskell):

- Generate input data based on a *strategy* (how should the data look?)
- Run the test case many times (say, 100), with different random generated data
- Check for properties / invariants that should hold for all input data
- If there is a failure, minimize the input data to a minimal failing example

```
Old buggy RPN calculator
```

```
— rpncalc/rpn _v1 .py -
class RPNCalculator:
                                       > 1
                                       Traceback (most recent call last):
    def evaluate(self, inp: str):
        if inp.isdigit():
            n = float(inp)
                                         File ..., in <module>
                                           rpn.run()
            self.stack.append(n)
        elif inp in "+-*/":
                                         File ..., in run
                                           self.evaluate(inp)
            b = self.stack.pop()
                                         File ..., in evaluate
            a = self.stack.pop()
            res = calc(a, b, inp)
                                           a = self.stack.pop()
                                       IndexError: pop from empty list
            self.stack.append(res)
            print(res)
```

Old buggy RPN calculator: Hypothesis test

```
-- plugins/test_hypothesis_rpncalc_v1.py
from hypothesis import given, strategies as st
from rpncalc.rpn_v1 import RPNCalculator

@given(st.text())
def test_random_strings(s: str):
    rpn = RPNCalculator()
    rpn.evaluate(s)
```

Old buggy RPN calculator: Hypothesis test

Old buggy RPN calculator: Hypothesis test

```
— plugins/test_hypothesis_rpncalc_v1.py
from hypothesis import given, strategies as st
from rpncalc.rpn_v1 import RPNCalculator
@given(st.text())
                                       > b = self.stack.pop()
def test_random_strings(s: str):
                                         IndexError: pop from empty list
    rpn = RPNCalculator()
                                         Falsifying example:
    rpn.evaluate(s)
                                              test_random_strings(s='')
                                      E
                                       >>> "" in "+-*/"
elif inp in "+-*/":
                                       True
```

```
Surely v2 is bug-free?
 — rpncalc/rpn _v2 .py
 class RPNCalculator:
      def evaluate(self, inp: str) -> None:
          if len(self.stack) < 2:
              print("Not enough operands")
              return
          b = self.stack.pop()
          a = self.stack.pop()
```

a = self.stack.pop()

Surely v2 is bug-free?

```
— rpncalc/rpn _v2 .py
class RPNCalculator:
    def evaluate(self, inp: str) -> None:
                                              test_random_strings PASSED
        if len(self.stack) < 2:
            print("Not enough operands")
            return
        b = self.stack.pop()
```

# Property-based testing with Hypothesis Surely v2 is bug-free?

```
— plugins/test_hypothesis_rpncalc_v2.py
@given(st.integers(), st.integers())
def test_operators(n1: int, n2: int):
    rpn = RPNCalculator(Config())
    rpn.evaluate(str(n1))
    rpn.evaluate(str(n2))
    rpn.evaluate("+")
    assert rpn.stack == [n1 + n2]
```

```
Surely v2 is bug-free?
 — plugins/test_hypothesis_rpncalc_v2.py
 @given(st.integers(), st.integers())
 def test_operators(n1: int, n2: int):
     rpn = RPNCalculator(Config())
     rpn.evaluate(str(n1))
     rpn.evaluate(str(n2))
     rpn.evaluate("+")
      assert rpn.stack == [n1 + n2]
     assert [1.055531162664984 8 e+16] == [1055531162664984 9]
     Falsifying example: test_operators(
         n1=0.
         n2=10555311626649849.
```

```
Surely v2 is bug-free?
```

```
— plugins/test_hypothesis_rpncalc_v2.py
@given(st.integers(), st.integers())
                                         > 10555311626649849
def test_operators(n1: int, n2: int):
                                         > 0
   rpn = RPNCalculator(Config())
   rpn.evaluate(str(n1))
                                         > +
   rpn.evaluate(str(n2))
                                         1.0555311626649848e+16
   rpn.evaluate("+")
    assert rpn.stack == [n1 + n2]
   assert [1.055531162664984 8 e+16] == [1055531162664984 9]
   Falsifying example: test_operators(
       n1=0,
       n2=10555311626649849.
```

```
Surely v2 is bug-free?
```

```
— plugins/test_hypothesis_rpncalc_v2.py
@given(st.integers(), st.integers())
                                         > 10555311626649849
def test_operators(n1: int, n2: int):
                                         > 0
   rpn = RPNCalculator(Config())
   rpn.evaluate(str(n1))
   rpn.evaluate(str(n2))
                                         1.0555311626649848e+16
   rpn.evaluate("+")
    assert rpn.stack == [n1 + n2]
   assert [1.055531162664984 8 e+16] == [1055531162664984 9]
   Falsifying example: test_operators(
                                          — rpncalc/rpn _v2 .py -
       n1=0,
       n2=10555311626649849,
                                          self.stack.append(float(inp))
```

More sophisticated strategies

```
@given(
    st.lists(
        st.one of(
            st.integers().map(str),
            st.floats().map(str),
            st.just("+"), st.just("-"),
            st.just("*"), st.just("/").
def test_usage(inputs: list[str]):
    rpn = RPNCalculator(Config())
    for inp in inputs:
        rpn.evaluate(inp)
```

More sophisticated strategies

```
@given(
    st.lists(
        st.one of(
            st.integers().map(str),
            st.floats().map(str),
            st.just("+"), st.just("-"),
            st.just("*"), st.just("/").
def test_usage(inputs: list[str]):
   rpn = RPNCalculator(Config())
    for inp in inputs:
        rpn.evaluate(inp)
```

```
Trying example: test_usage(
  inputs=[
    '-2.2250738585e-313',
    '/', '-10', '110', '+', '+'])
Not enough operands
100.0
100.0
```

## Property-based testing with Hypothesis More sophisticated strategies

```
@given(
    st.lists(
        st.one of(
            st.integers().map(str),
            st.floats().map(str),
            st.just("+"), st.just("-"),
            st.just("*"), st.just("/"),
def test_usage(inputs: list[str]):
    rpn = RPNCalculator(Config())
    for inp in inputs:
        rpn.evaluate(inp)
```

```
Trying example: test_usage(
  inputs=[
    '-2.2250738585e-313'.
    '/'. '-10'. '110'. '+'. '+'])
Not enough operands
100.0
100.0
Trying example: test_usage(
  inputs=['+', '+', '-', '*'],
Not enough operands
Not enough operands
Not enough operands
Not enough operands
```

Another example: Run length encoding

Original:

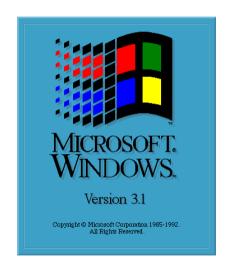
PPPPPYYYPPPPPPYYYY

Compressed:

5P3Y7P4Y

#### Used in:

- Fax machines
- Windows 3.1 startup screen
- . . .



# Property-based testing with Hypothesis Decoding

```
input: [(5, "P"), (3, "Y"), (7, "P"), (4, "Y")]
output: "PPPPPYYYPPPPPPYYYY"
— plugins/test_hypothesis.py
def decode(lst: list[tuple[int, str]]) -> str:
    for count, character in 1st:
        s += count * character
    return s
```

```
def encode(input_string: str) -> list[tuple[int, str]]:
    count = 1; prev = ""; lst = []
                                               input:
    for character in input_string:
                                               "PPPPPYYYPPPPPPPYYYY"
        if character != prev:
            if prev:
                                               output:
                entry = (count, prev)
                lst.append(entry)
                                                  (5, "P"),
            count = 1
                                                  (3. "Y").
            prev = character
                                                  (7, "P"),
        else:
                                                  (4. "Y").
            count += 1
    entry = (count, character)
    lst.append(entry)
    return 1st
```

```
def encode(input_string: str) -> list[tuple[int, str]]:
    count = 1; prev = ""; lst = []
                                               input:
    for character in input_string:
                                               "PPPPPYYYPPPPPPPYYYY"
        if character != prev:
            if prev:
                                               output:
                 entry = (count, prev)
                lst.append(entry)
                                                  (5, "P").
            count = 1
                                                  (3. "Y").
            prev = character
                                                  (7, "P"),
        else:
                                                  (4. "Y").
            count += 1
    entry = (count, character)
    lst.append(entry)
    return 1st
```

```
def encode(input_string: str) -> list[tuple[int, str]]:
    count = 1; prev = ""; lst = []
                                               input:
    for character in input_string:
                                               "PPPPPYYYPPPPPPPYYYY"
        if character != prev:
            if prev:
                                               output:
                 entry = (count, prev)
                lst.append(entry)
                                                  (5, "P").
            count = 1
                                                  (3. "Y").
            prev = character
                                                  (7, "P"),
        else:
                                                  (4. "Y").
            count += 1
    entry = (count, character)
    lst.append(entry)
    return 1st
```

```
def encode(input_string: str) -> list[tuple[int, str]]:
    count = 1; prev = ""; lst = []
                                               input:
    for character in input_string:
                                               "PPPPPYYYPPPPPPPYYYY"
        if character != prev:
            if prev:
                                               output:
                entry = (count, prev)
                lst.append(entry)
                                                  (5, "P").
            count = 1
                                                  (3. "Y").
            prev = character
                                                  (7, "P"),
        else:
                                                  (4. "Y").
            count += 1
    entry = (count, character)
    lst.append(entry)
    return 1st
```

```
def encode(input_string: str) -> list[tuple[int, str]]:
    count = 1; prev = ""; lst = []
                                               input:
    for character in input_string:
                                               "PPPPPYYYPPPPPPPYYYY"
        if character != prev:
            if prev:
                                               output:
                entry = (count, prev)
                lst.append(entry)
                                                  (5, "P").
            count = 1
                                                  (3. "Y").
            prev = character
                                                  (7, "P"),
        else:
                                                  (4. "Y").
            count += 1
    entry = (count, character)
    lst.append(entry)
    return 1st
```

```
def encode(input_string: str) -> list[tuple[int, str]]:
    count = 1; prev = ""; lst = []
                                               input:
    for character in input_string:
                                               "PPPPPYYYPPPPPPPYYYY"
        if character != prev:
            if prev:
                                               output:
                entry = (count, prev)
                lst.append(entry)
                                                  (5, "P").
            count = 1
                                                  (3. "Y").
            prev = character
                                                  (7, "P"),
        else:
                                                  (4. "Y").
            count += 1
    entry = (count, character)
    lst.append(entry)
    return 1st
```

# Property-based testing with Hypothesis Manual test

Can we think of all corner-cases?

Maybe it's easier to think of *invariants* (certain properties that always hold true)?

Test with Hypothesis

```
@given(text())

def test_decode_inverts_encode(s: str):
    assert decode(encode(s)) == s
```

Result

#### UnboundLocalError:

local variable 'character'
referenced before assignment

### Falsifying example:

test\_decode\_inverts\_encode(s='')

### Demo plugins/test\_hypothesis.py:

- Fix the issue, e.g. by adding if not input\_string: return []
   at the beginning of the encode-function.
- Re-run the test, it should pass.
   Run with --hypothesis-verbosity=verbose and -s to observe examples.
- Deliberately introduce a more subtle issue, e.g. break resetting the count by commenting out the second count = 1.
- Re-run the test you should get a minimal example which fails the test.

Another possibility

```
def encode itertools(input string: str) -> list[tuple[int, str]]:
    return [
        (len(list(group)), character)
        for character, group
        in itertools.groupby(input_string)
@given(text())
def test_alternative_implementation(s: str):
    assert encode_itertools(s) == encode(s)
```

What it can do

- Generate data based on types/strategies
- Combine/filter various strategies to generate more sophisticated data
- Generate names, etc. via fakefactory
- Generate data matching a Django model
- Generate data matching a grammar
- Generate test code (rather than data) based on a state machine

## Upcoming events

March 4th to 6th, 2025
 Python Academy
 (python-academy.com):
 Professional Testing with Python
 Leipzig, Germany & Remote

### Custom training / coaching:

- Python
- pytest
- GUI programming with Qt
- Best Practices (packaging, linting, etc.)
- Git
- . . .

Remote or on-site florian@bruhin.software https://bruhin.software/





# Feedback and questions

- Service Properties
  P
- ☑ florian@bruhin.software
- 6 bruhin.software
- X @the compiler
- in linkedin.com/in/florian-bruhin

Copyright 2015 - 2024 Florian Bruhin





CC BY 4.0





