



Czy na MuZero można grać w Doom'a?

Paulina Brzęcka Marek Borzyszkowski

19 stycznia 2025



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1].



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1]. Dziś dowiedzie się:

1. na czym polega testowanie na podstawie właściwości.
2. Czym różni się ono od klasycznego podejścia do testowania.
3. Jakie są strategie testowania na podstawie właściwości.
4. Jak przy wykorzystaniu konceptu QuickCheck znaleźć wartości początkowe nieprzechodzące testy.



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1].



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1]. Dziś dowiedzie się:

1. na czym polega testowanie na podstawie właściwości.
2. Czym różni się ono od klasycznego podejścia do testowania.
3. Jakie są strategie testowania na podstawie właściwości.
4. Jak przy wykorzystaniu konceptu QuickCheck znaleźć wartości początkowe nieprzechodzące testy.



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1].



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1]. Dziś dowiedzie się:

1. na czym polega testowanie na podstawie właściwości.
2. Czym różni się ono od klasycznego podejścia do testowania.
3. Jakie są strategie testowania na podstawie właściwości.
4. Jak przy wykorzystaniu konceptu QuickCheck znaleźć wartości początkowe nieprzechodzące testy.



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1].



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1]. Dziś dowiedzie się:

1. na czym polega testowanie na podstawie właściwości.
2. Czym różni się ono od klasycznego podejścia do testowania.
3. Jakie są strategie testowania na podstawie właściwości.
4. Jak przy wykorzystaniu konceptu QuickCheck znaleźć wartości początkowe nieprzechodzące testy.



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1].



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1]. Dziś dowiedzie się:

1. na czym polega testowanie na podstawie właściwości.
2. Czym różni się ono od klasycznego podejścia do testowania.
3. Jakie są strategie testowania na podstawie właściwości.
4. Jak przy wykorzystaniu konceptu QuickCheck znaleźć wartości początkowe nieprzechodzące testy.



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1].



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1]. Dziś dowiedzie się:

1. na czym polega testowanie na podstawie właściwości.
2. Czym różni się ono od klasycznego podejścia do testowania.
3. Jakie są strategie testowania na podstawie właściwości.
4. Jak przy wykorzystaniu konceptu QuickCheck znaleźć wartości początkowe nieprzechodzące testy.



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1].



Istnieje wiele koncepcji testowania oprogramowania, jedną z nich jest testowanie na podstawie właściwości [1]. Dziś dowiedzie się:

1. na czym polega testowanie na podstawie właściwości.
2. Czym różni się ono od klasycznego podejścia do testowania.
3. Jakie są strategie testowania na podstawie właściwości.
4. Jak przy wykorzystaniu konceptu QuickCheck znaleźć wartości początkowe nieprzechodzące testy.



Testowanie na podstawie właściwości nie jest proste w zastosowaniu, przynajmniej przy pierwszej próbie. Istnieją jednak pewne schematy wskazujące drogę, jak stworzyć takie testy.

1. Różne ścieżki, ten sam wynik



Testowanie na podstawie właściwości nie jest proste w zastosowaniu, przynajmniej przy pierwszej próbie. Istnieją jednak pewne schematy wskazujące drogę, jak stworzyć takie testy.

1. Różne ścieżki, ten sam wynik
2. Tam i z powrotem



Testowanie na podstawie właściwości nie jest proste w zastosowaniu, przynajmniej przy pierwszej próbie. Istnieją jednak pewne schematy wskazujące drogę, jak stworzyć takie testy.

1. Różne ścieżki, ten sam wynik
2. Tam i z powrotem
3. Są rzeczy niezmiennie



Testowanie na podstawie właściwości nie jest proste w zastosowaniu, przynajmniej przy pierwszej próbie. Istnieją jednak pewne schematy wskazujące drogę, jak stworzyć takie testy.

1. Różne ścieżki, ten sam wynik
2. Tam i z powrotem
3. Są rzeczy niezmiennie
4. Z czasem rzeczy przestają się zmieniać



Testowanie na podstawie właściwości nie jest proste w zastosowaniu, przynajmniej przy pierwszej próbie. Istnieją jednak pewne schematy wskazujące drogę, jak stworzyć takie testy.

1. Różne ścieżki, ten sam wynik
2. Tam i z powrotem
3. Są rzeczy niezmiennie
4. Z czasem rzeczy przestają się zmieniać
5. Dziel i rządź



Testowanie na podstawie właściwości nie jest proste w zastosowaniu, przynajmniej przy pierwszej próbie. Istnieją jednak pewne schematy wskazujące drogę, jak stworzyć takie testy.

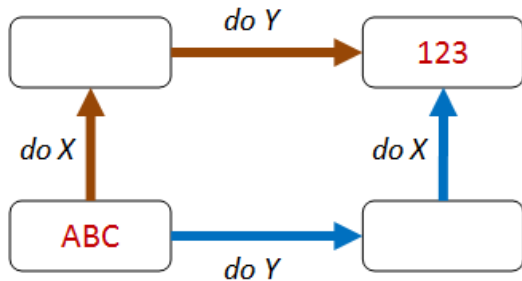
1. Różne ścieżki, ten sam wynik
2. Tam i z powrotem
3. Są rzeczy niezmiennie
4. Z czasem rzeczy przestają się zmieniać
5. Dziel i rządź
6. Łatwiej zweryfikować niż zaimplementować



Testowanie na podstawie właściwości nie jest proste w zastosowaniu, przynajmniej przy pierwszej próbie. Istnieją jednak pewne schematy wskazujące drogę, jak stworzyć takie testy.

1. Różne ścieżki, ten sam wynik
2. Tam i z powrotem
3. Są rzeczy niezmiennie
4. Z czasem rzeczy przestają się zmieniać
5. Dziel i rządź
6. Łatwiej zweryfikować niż zaimplementować
7. Testowanie z wyrocznią

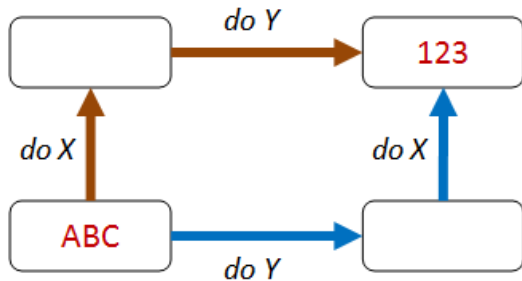
Jedną z podstawowych strategii skorzystanie z komutatywności niektórych operacji. Można to zrobić poprzez wykonanie operacji w różnej kolejności.



Rysunek: Strategia - komutatywność

Jedną z podstawowych strategii skorzystanie z komutatywności niektórych operacji. Można to zrobić poprzez wykonanie operacji w różnej kolejności.

Przykładem takiej strategii może być komutatywność dodawania, gdzie wykorzystano $\text{add } x \ y$, jak i odwrotność tej operacji $\text{add } y \ x$.

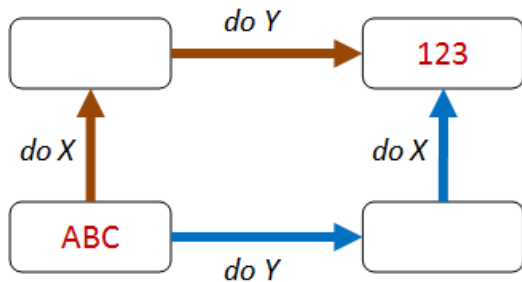


Rysunek: Strategia - komutatywność

Jedną z podstawowych strategii skorzystanie z komutatywności niektórych operacji. Można to zrobić poprzez wykonanie operacji w różnej kolejności.

Przykładem takiej strategii może być komutatywność dodawania, gdzie wykorzystano $\text{add } x \ y$, jak i odwrotność tej operacji $\text{add } y \ x$.

Innym przykładem jest test metody sort danej listy. Wykonanie sortowania, a następnie dodanie do każdego elementu listy 1 powinno dać taki sam efekt jak dodanie 1 do każdego z elementów listy, a następnie jej posortowanie.

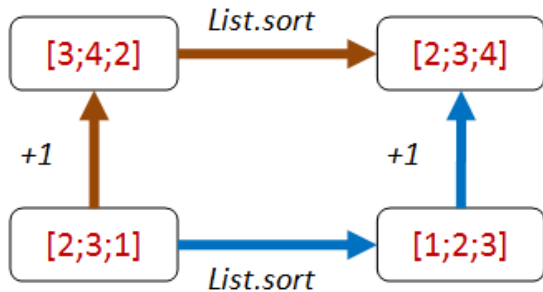


Rysunek: Strategia - komutatywność

Jedną z podstawowych strategii skorzystanie z komutatywności niektórych operacji. Można to zrobić poprzez wykonanie operacji w różnej kolejności.

Przykładem takiej strategii może być komutatywność dodawania, gdzie wykorzystano $\text{add } x \ y$, jak i odwrotność tej operacji $\text{add } y \ x$.

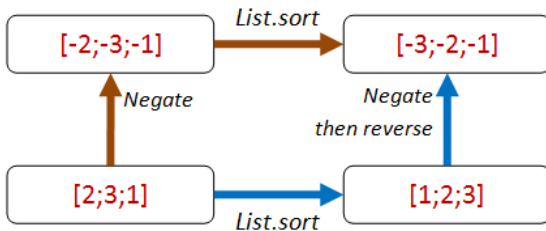
Innym przykładem jest test metody `sort` danej listy. Wykonanie sortowania, a następnie dodanie do każdego elementu listy 1 powinno dać taki sam efekt jak dodanie 1 do każdego z elementów listy, a następnie jej posortowanie.



Rysunek: Strategia - komutatywność



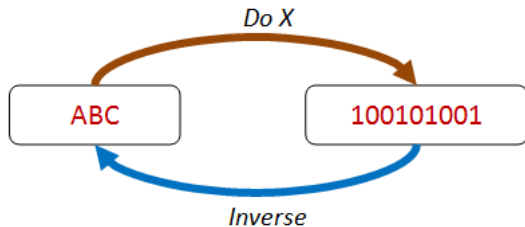
```
let addThenSort_eq_sortThenAdd sortFn aList =  
  let add1 x = x + 1  
  
  let result1 = aList |> sortFn |> List.map add1  
  let result2 = aList |> List.map add1 |> sortFn  
  result1 = result2
```



Rysunek: Strategia - komutatywność

Przykładem takiego testu mogą być przeciwne operacje matematyczne jak:

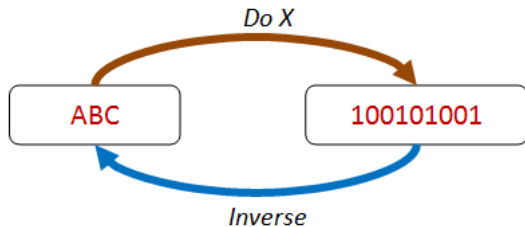
- dodawanie/odejmowanie



Rysunek: Strategia - inwersja

Przykładem takiego testu mogą być przeciwne operacje matematyczne jak:

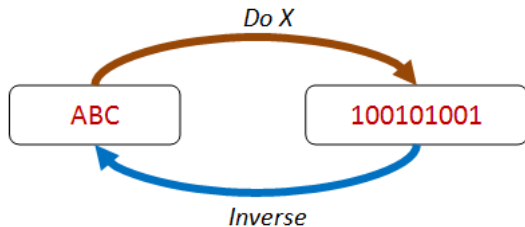
- dodawanie/odejmowanie
- mnożenie/dzielenie



Rysunek: Strategia - inwersja

Przykładem takiego testu mogą być przeciwne operacje matematyczne jak:

- dodawanie/odejmowanie
- mnożenie/dzielenie
- potęga/logarytm.



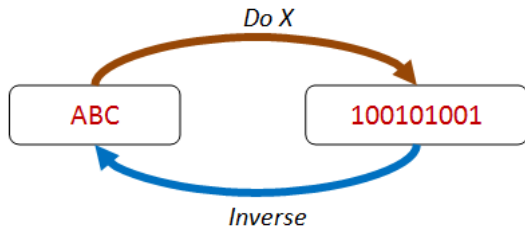
Rysunek: Strategia - inwersja

Przykładem takiego testu mogą być przeciwne operacje matematyczne jak:

- dodawanie/odejmowanie
- mnożenie/dzielenie
- potęga/logarytm.

Innymi przykładami są operacje niekoniecznie matematyczne:

- serializacja/deserializacja



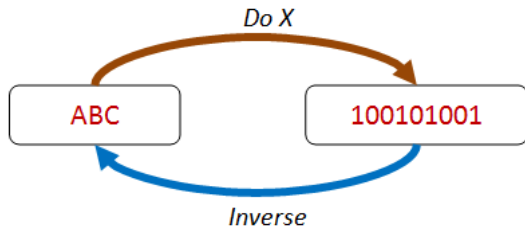
Rysunek: Strategia - inwersja

Przykładem takiego testu mogą być przeciwne operacje matematyczne jak:

- dodawanie/odejmowanie
- mnożenie/dzielenie
- potęga/logarytm.

Innymi przykładami są operacje niekoniecznie matematyczne:

- serializacja/deserializacja
- zapis/odczyt z pliku



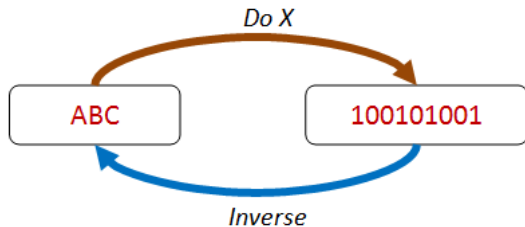
Rysunek: Strategia - inwersja

Przykładem takiego testu mogą być przeciwne operacje matematyczne jak:

- dodawanie/odejmowanie
- mnożenie/dzielenie
- potęga/logarytm.

Innymi przykładami są operacje niekoniecznie matematyczne:

- serializacja/deserializacja
- zapis/odczyt z pliku
- wstaw/sprawdź czy zawiera.



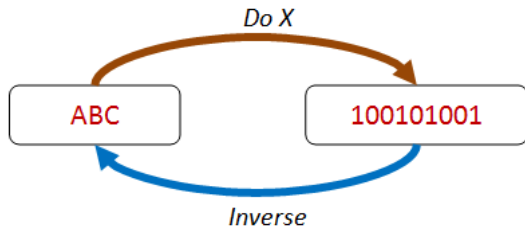
Rysunek: Strategia - inwersja

Przykładem takiego testu mogą być przeciwne operacje matematyczne jak:

- dodawanie/odejmowanie
- mnożenie/dzielenie
- potęga/logarytm.

Innymi przykładami są operacje niekoniecznie matematyczne:

- serializacja/deserializacja
- zapis/odczyt z pliku
- wstaw/sprawdź czy zawiera.
- odwrócenie listy/odwrócenie listy



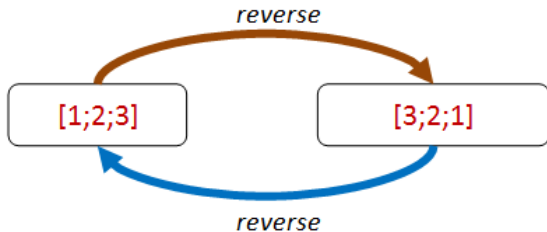
Rysunek: Strategia - inwersja

Przykładem takiego testu mogą być przeciwne operacje matematyczne jak:

- dodawanie/odejmowanie
- mnożenie/dzielenie
- potęga/logarytm.

Innymi przykładami są operacje niekoniecznie matematyczne:

- serializacja/deserializacja
- zapis/odczyt z pliku
- wstaw/sprawdź czy zawiera.
- odwrócenie listy/odwrócenie listy



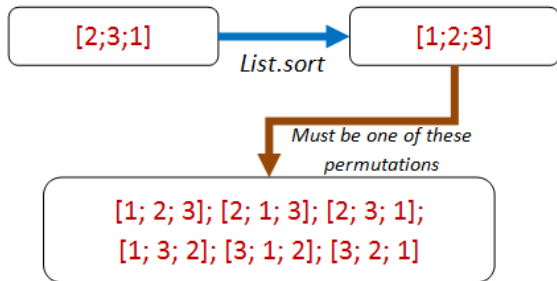
Rysunek: Strategia - inwersja

Czasami testowana funkcja przetwarzając dane zachowuje część ich właściwości. Chociażby funkcje `sort` lub `map` wykonane na liście n elementów, zwracają odpowiednio zmodyfikowaną listę n elementową.



Rysunek: Strategia - niezmiennosc

Czasami testowana funkcja przetwarzając dane zachowuje część ich właściwości. Chociażby funkcje sort lub map wykonane na liście n elementów, zwracają odpowiednio zmodyfikowaną listę n elementową.



Rysunek: Strategia - niezmiennosc



Strategie - Z czasem rzeczy przestają się zmieniać

Inną właściwością funkcji może być niezmiennosc wyniku funkcji po ponownym jej zaaplikowaniu. Innymi słowy, wykonanie funkcji 2 razy daje taki sam efekt, jak jednokrotne jej zaaplikowanie.



Rysunek: Strategia - idempotentność



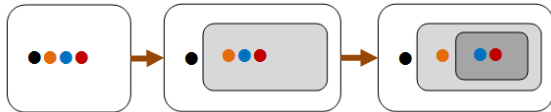
Strategie - Z czasem rzeczy przestają się zmieniać

Inną właściwością funkcji może być niezmiennosc wyniku funkcji po ponownym jej zaaplikowaniu. Innymi słowy, wykonanie funkcji 2 razy daje taki sam efekt, jak jednokrotne jej zaaplikowanie. Przykładami takich operacji, dla których taki typ testu miałby zastosowanie to metoda `distinct` wykonana na danej liście, lub wykonanie `update` na danej bazie danych.



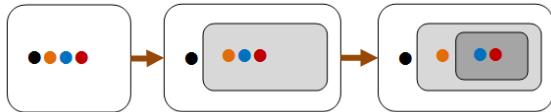
Rysunek: Strategia - idempotentność

Istnieją sposoby na testowanie na podstawie właściwości jest wykorzystanie rekursywności struktur przekazywanych do funkcji, takich jak listy, drzewa.



Rysunek: Strategia - rekursywna

Istnieją sposoby na testowanie na podstawie właściwości jest wykorzystanie rekursywności struktur przekazywanych do funkcji, takich jak listy, drzewa. Przykładem może być sprawdzenie za pomocą tej metody funkcji sort.



Rysunek: Strategia - rekursywna

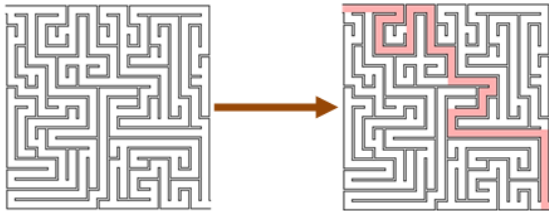


```
let rec firstLessThanSecond_andTailIsSorted sortFn
(aList:int list) =
    let sortedList = aList |> sortFn
    match sortedList with
    | [] -> true
    | [first] -> true
    | [first;second] -> first <= second
    | first::second::rest->
        first <= second &&
        let tail = second::rest
        // check that tail is sorted
        firstLessThanSecond_andTailIsSorted sortFn tail
```



Strategie - Łatwiej zweryfikować niż zaimplementować

Niekiedy testowana funkcja jest skomplikowana, ale jej rezultat da się łatwo sprawdzić. Przykładem może być funkcja wyszukująca wyjście z labiryntu, gdzie sam algorytm wyszukiwania odpowiedniej ścieżki jest skomplikowany, natomiast samo sprawdzenie, czy ścieżka dobrze prowadzi do wyjścia można w łatwy sposób zweryfikować.

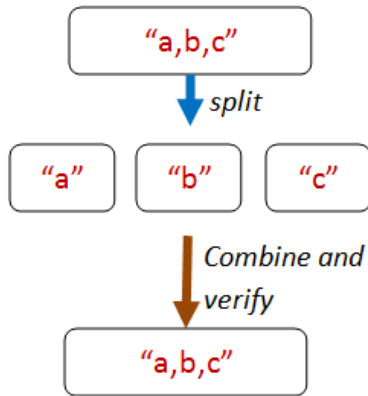


Rysunek: Strategia - łatwe sprawdzenie



Strategie - Łatwiej zweryfikować niż zaimplementować

Niekiedy testowana funkcja jest skomplikowana, ale jej rezultat da się łatwo sprawdzić. Przykładem może być funkcja wyszukująca wyjście z labiryntu, gdzie sam algorytm wyszukiwania odpowiedniej ścieżki jest skomplikowany, natomiast samo sprawdzenie, czy ścieżka dobrze prowadzi do wyjścia można w łatwy sposób zweryfikować.





Strategie - Łatwiej zweryfikować niż zaimplementować

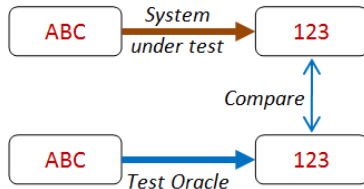
```
let concatElementsOfSplitString_eq_originalString (strings:string list)
  // make a string
  let inputString = strings |> String.concat ","

  // use the 'split' function on the inputString
  let tokens = stringSplit inputString

  // build a string from the tokens
  let recombinedString = tokens |> String.concat ","

  // compare the result with the original
  inputString = recombinedString
```

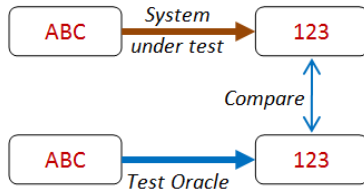
Zdaża się, że funkcjonalność została już napisana i trzeba ją zrefactorować, przepisać, napisać od nowa. Warto wtedy wykorzystać wartości zwracane przez oryginalnie zaimplementowany algorytm jako pewną wartość wyniku, pewnego rodzaju wyrocznię, uznając go jako prawdę.



Rysunek: Strategia - wyrocznia

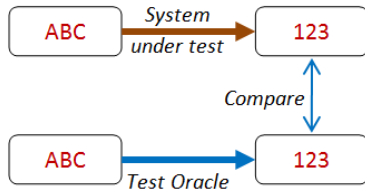


Zdaje się, że funkcjonalność została już napisana i trzeba ją zrefactorować, przepisać, napisać od nowa. Warto wtedy wykorzystać wartości zwracane przez oryginalnie zaimplementowany algorytm jako pewną wartość wyniku, pewnego rodzaju wyrocznię, uznając go jako prawdę. W taki sposób można sprawdzić, czy nowa funkcja w pewnym stopniu pokrywa się ze starą funkcją. Czasami też istnieje wiele algorytmów doprowadzających do tego samego wyniku, mające różne złożoności, czy też działające równolegle.



Rysunek: Strategia - wyrocznia

Zdaje się, że funkcjonalność została już napisana i trzeba ją zrefactorować, przepisać, napisać od nowa. Warto wtedy wykorzystać wartości zwracane przez oryginalnie zaimplementowany algorytm jako pewną wartość wyniku, pewnego rodzaju wyrocznię, uznając go jako prawdę. W taki sposób można sprawdzić, czy nowa funkcja w pewnym stopniu pokrywa się ze starą funkcją. Czasami też istnieje wiele algorytmów doprowadzających do tego samego wyniku, mające różne złożoności, czy też działające równolegle. Można wykozystać wtedy najprostrzy algorytm jako wyrocznię, ze względu na najmniejsze prawdopodobieństwo napisania takiego algorytmu z błędem. Następnie, przy wykorzystaniu wyroczni, stworzyć bardziej skomplikowany (często efektywniejszy) algorytm.



Rysunek: Strategia - wyrocznia

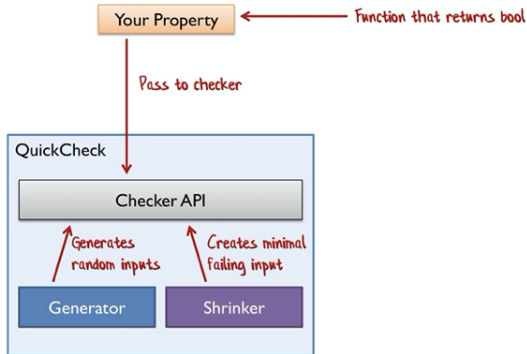


QuickCheck został stworzony w języku Haskell jako pierwsze narzędzie wspierające testowanie oparte na właściwościach. Jest inspiracją dla bibliotek w innych językach, takich jak FsCheck dla .NET.



- Automatyczne generowanie danych testowych.
- Sprawdzanie, czy zdefiniowane właściwości funkcji są spełniane dla wielu losowych przypadków.
- Redukcja (*shrinking*) danych wejściowych w przypadku błędu.

1. **Checker API** wykrywa typ wejścia funkcji.
2. Wywoływany jest generator odpowiedniego typu.
3. Następuje generowanie przypadków testowych.
4. Przypadki testowe są przekazywane do testowanej właściwości.



Rysunek: Schemat działania QuickCheck



- **Check.Quick** – uruchamia szybki test, sprawdzając właściwość dla domyślnej liczby losowych przypadków (np. 100).
- **Check.Verbose** – działa jak Check.Quick, ale wyświetla więcej szczegółowych informacji o danych testowych.
- **Generowanie danych wejściowych** – FsCheck wspiera generowanie danych dla typów prostych (np. `int`, `float`, `string`) i bardziej złożonych struktur, takich jak listy czy rekordy.



FsCheck umożliwia definiowanie właściwości w formie funkcji logicznych. Właściwości opisują, jakie warunki zawsze muszą być spełnione przez funkcję.

- Odwrócenie listy dwukrotnie powinno zwrócić pierwotną listę.

```
let reverseProperty (xs: int list) =  
    List.rev (List.rev xs) = xs
```

```
Check.Quick reverseProperty
```



FsCheck pozwala na tworzenie własnych generatorów danych wybranego typu.

- Tworzenie generatora.
- Generowanie danych testowych dla różnych struktur.
- Wykorzystanie arbitrażu (Arb) do definiowania generatorów.



```
let intGenerator = Arb.generate<int>
Gen.sample 100 3 intGenerator // [-37; 24; -62]
type PositiveInt = PositiveInt of int
let positiveIntGen =
    Gen.suchThat ((<) 0) Arb.generate<int>
Arb.register<PositiveInt>(positiveIntGen)
type Generators =
    static member PositiveInt() =
        Arb.fromGen positiveIntGen
Arb.register<Generators>() // rejestracja generatora
let intListGenerator = Arb.generate<int list>
Gen.sample 5 10 intListGenerator
// [ []; []; [-4]; [0; 3; -1; 2]; [1]; [1]; []; [0; 1; -2]; [] ]
let stringGenerator = Arb.generate<string>
Gen.sample 10 3 stringGenerator // [""; "eiX$a^"; "U%OIka&r"]
```



```
type Point = {x:int; y:int; color: Color}
let pointGenerator = Arb.generate<Point>
Gen.sample 50 10 pointGenerator
(*
{x = -8; y = 12; color = Green -4;};
{x = 28; y = -31; color = Green -6;};
{x = 11; y = 27; color = Red;};
{x = -2; y = -13; color = Red;};
{x = 6; y = 12; color = Red;};
// itd
*)
```



Proces redukcji danych wejściowych, który minimalizuje przypadki testowe prowadzące do błędu.

- Redukcja liczb do coraz mniejszych wartości.
- Redukcja ciągów znaków.

```
Arb.shrink "abcd" |> Seq.toList  
//["bcd"; "acd"; "abd"; "abc"; "abca"; "abcb"; "abcc"; "abad"]
```



```
let isSmallerThan80 x = x < 80
isSmallerThan80 100 // false, so start shrinking

Arb.shrink 100 |> Seq.toList// [0; 50; 75; 83; 94; 97; 99]
isSmallerThan80 0 // true
isSmallerThan80 50 // true
isSmallerThan80 75 // true
isSmallerThan80 83 // false, so shrink again

Arb.shrink 83 |> Seq.toList// [0; 44; 77; 79; 80; 81]
isSmallerThan80 0 // true
isSmallerThan80 44 // true
isSmallerThan80 77 // true
isSmallerThan80 79 // true
isSmallerThan80 80 // false <- najmniejsza porażka
// wynik: Falsifiable, after 10 tests (2 shrinks)
```



```
let config = {  
    Config.Quick with  
        MaxTest = 1000  
}  
Check.One(config, isSmallerThan80 )  
// result: Ok, passed 1000 tests. (a nie powinno :)  
  
let config = {  
    Config.Quick with  
        MaxTest = 10000  
}  
Check.One(config, isSmallerThan80 )  
// result: Falsifiable, after 8660 tests (1 shrink):  
//      80
```



```
let precondition x y =  
  (x,y) <> (0,0)  
  && (x,y) <> (2,2)
```

```
let additionIsNotMultiplication_withPreCondition x y =  
  precondition x y ==> additionIsNotMultiplication x y  
// Ok, passed 100 tests.
```



```
type AdditionSpecification =  
  static member ``Commutative`` x y =  
    commutativeProperty x y  
  static member ``Associative`` x y z =  
    associativeProperty x y z  
  static member ``Left Identity`` x =  
    leftIdentityProperty x  
  static member ``Right Identity`` x =  
    rightIdentityProperty x
```

```
Check.QuickAll<AdditionSpecification>()  
--- Checking AdditionSpecification ---  
AdditionSpecification.Commutative-Ok, passed 100 tests.  
AdditionSpecification.Associative-Ok, passed 100 tests.  
AdditionSpecification.Left Identity-Ok, passed 100 tests.  
AdditionSpecification.Right Identity-Ok, passed 100 tests.
```



TU JEST PODSUMOWANIE



Pytania?



- [1] S. Wlaschin, “The ”Property Based Testing” series”,, s. 6, grud. 2014. adr.:
<https://fsharpforfunandprofit.com/posts/property-based-testing/>.



Chcielibyśmy podziękować Panu dr. inż. Janowi Cychnerskiemu za stworzenie i udostępnienie stylu *pg-beamer*, co zostało wykorzystane do stworzenia tej prezentacji.
<https://github.com/jachoo/pg-beamer>



Dziękujemy za uwagę!



**POLITECHNIKA
GDAŃSKA**