Building test check Generators

Gary Fredericks

Who Am I?

- I am Gary Fredericks
- I live in Chicago
- I work (with Clojure) at DRW
- I sometimes work on improving test.check

• test.check

- test.check
 - Created 4 years ago by Reid Draper
 - Clojure library for property-based-testing
 - Haskell's QuickCheck
 - Inputs described by writing generators

- test.check
 - Created 4 years ago by Reid Draper
 - Clojure library for property-based-testing
 - Haskell's QuickCheck
 - Inputs described by writing generators
- clojure.spec

- test.check
 - Created 4 years ago by Reid Draper
 - Clojure library for property-based-testing
 - Haskell's QuickCheck
 - Inputs described by writing generators
- clojure.spec
 - Announced 18 months ago
 - Creates generates for basic specs
 - User-supplied generators for complex specs

Roadmap

- Generators, in General
- Building
- Fine-Tuning

Generators, in General

Clojure.spec (.spec[.alpha] Making specs

C.spec.test[.alpha] instrument, check, etc.

C.spec.gen[alpha]

Dynaload!

test.check

c.t.check c.t.check.properties

Assembling/running tests

c.t.check.clojure-test

c.t.check.generators

Building generators

Generators, in General

What's in the box?

Basic Data Generators

```
1
      (def generate-some-great-data
 2
3
4
5
6
7
        (gen/hash-map
         :a-boolean
                                  gen/boolean
         :some-small-integers
                                 (gen/vector gen/nat)
         :a-large-integer
                                  gen/large-integer
         :a-double
                                  gen/double
         :a-color
                                  (gen/elements [:red :green :blue])
 8
                                  gen/uuid
         :a-uuid
 9
         :a-string-and-a-keyword (gen/tuple gen/string
10
                                              gen/keyword)))
11
12
      (gen/generate generate-some-great-data 10)
13
      =>
14
      {:a-boolean
                                false.
15
                            [2 5 5 5 10 5 0],
       :some-small-integers
16
       :a-large-integer
                                -6.
17
       :a-double
                                0.47607421875,
18
       :a-color
                                :green
19
                                #uuid "a06e2893-6fcc-4b42-8e2f-ba5da58202ac",
       :a-mid
20
       :a-string-and-a-keyword ["o") :eA:5C*:02]}
```

Combinators!

```
(gen/tuple g1 g2 ...)
(gen/fmap (fn [x] x') g)
(gen/bind g (fn [x] g'))
(gen/such-that pred g)
(gen/frequency [[w1 g1] [w2 g2] ...])
(gen/one-of [g1 g2 ...])
```

() -> Data

```
1  (defn gen-fav-number-assertion
2  []
3   (let [x (rand-int 10)]
4      (str "My favorite number is " x)))
5   
6   (gen-fav-number-assertion)
7  => "My favorite number is 9"
```

Size -> Data

```
1  (defn gen-fav-number-assertion
2    [size]
3    (let [x (rand-int size)]
4         (str "My favorite number is " x)))
5    (gen-fav-number-assertion 100)
7    => "My favorite number is 91"
```

Size -> (Data, Shrinks Data)

```
(defn gen-fav-number-assertion
  size
  (let [x (rand-int size)]
    [(str "My favorite number is " x)
     ;; some sort of recursively lazy
     ;; expression that generates a
     ;; lazy tree of smaller strings
    1))
(gen-fav-number-assertion 100)
=> ["My favorite number is 91" (...)]
```

2

3

4

5

6

7

8

9 10

Rand Double -> Size -> (Data, Shrinks Data)

```
(defn gen-fav-number-assertion
  rng size
  (let [x (-> rng (rand/rand-double) (* size) (long))]
    [(str "My favorite number is " x)
     ;; some sort of recursively lazy
     ;; expression that generates a
     ;; lazy tree of smaller strings
    1))
(gen-fav-number-assertion (rand/make-random 42) 100)
=> ["My favorite number is 91" (...)]
```

1

2

3

4

5

6

7

8

9 10

Composition

1

2

3

4

5

6

8

```
(defn collection-of
  "Returns a generator of a collection
 with elements generated from the
 supplied generator."
  gen
  (fn [rng size]
    ;; and now for the tricky bit
     ;; um
    1))
```

Value Proposition

```
clojure.core/rand-int
vs
gen/large-integer
```

- Abstract sizing/growth
- Shrinking
- Functional Determinism

Dev Tools

```
1 (gen/sample g <num-samples=10>)
2 => (data0 data1 data2 ...)
3 
4 (gen/generate g <size=30>)
5 => data
```

Universality & Purity

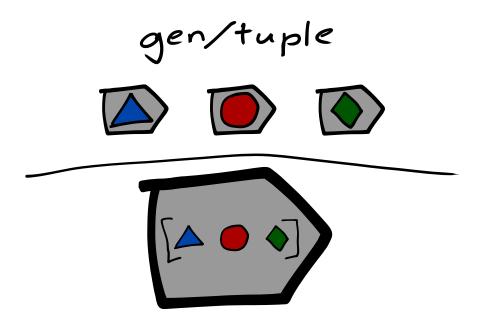
- gen/tuple
- gen/one-of
- gen/frequency
- gen/such-that
- gen/fmap
- gen/bind

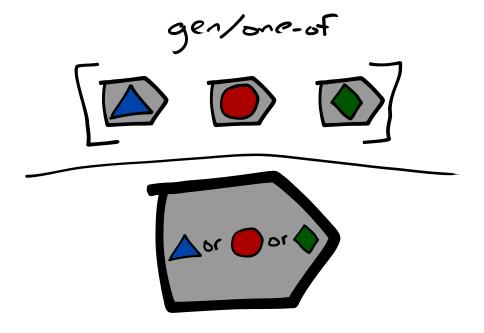
Similarities with FP:

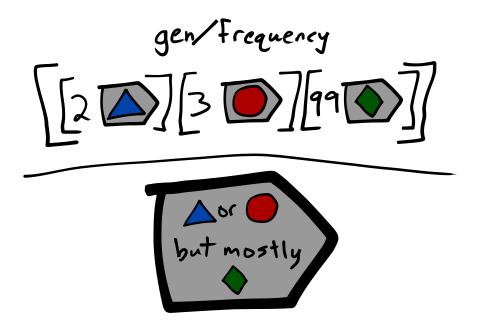
- Universality
 - The built-in generators can generate just about anything
 - But understanding how takes practice
- Circumventability
 - e.g.,
 - rand, etc.
 - gen/sample, gen/generate
 - using today's date, DB records
 - Undermines the value proposition
 - Abstract sizing/growth
 - Shrinking
 - Functional Determinism

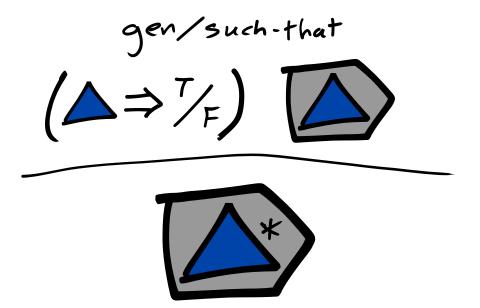
Combinators, this time in detail

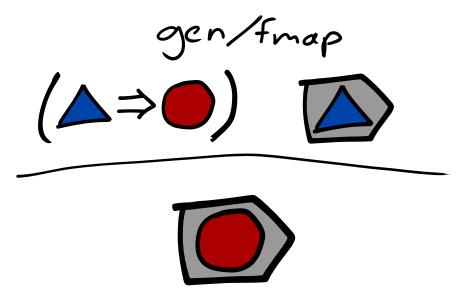
```
(gen/tuple g1 g2 ...)
(gen/fmap (fn [x] x') g)
(gen/bind g (fn [x] g'))
(gen/such-that pred g)
(gen/frequency [[w1 g1] [w2 g2] ...])
(gen/one-of [g1 g2 ...])
```

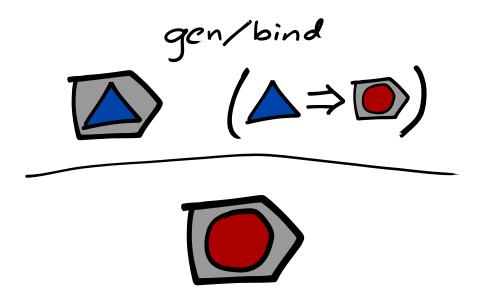


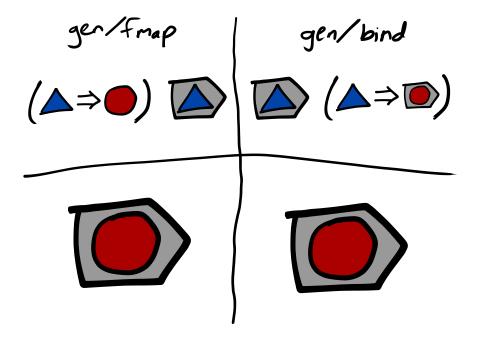












Example:

Generate a non-empty collection and a random element from that collection.

2

3

4

5

6

```
1
     (def gen-collection-and-element
 2
       (gen/fmap (fn [coll]
3
                    :: pick a random element, then use
4
                   ;; gen/fmap to wrap it up with the
 5
                   :: collection
 6
                   (gen/fmap (fn [x] [coll x])
 7
                              (gen/elements coll)))
 8
                 (gen/not-empty
 9
                  (gen/vector gen/large-integer))))
10
11
     (gen/generate gen-collection-and-element)
12
     #clojure.test.check.generators.Generator{:gen #function[clo
13
     ;; we just generated a generator, not data
```

```
(def gen-collection-and-element
  (gen/bind (gen/not-empty
             (gen/vector gen/large-integer))
            (fn [coll]
              ;; pick a random element, then use
              ;; gen/fmap to wrap it up with the
              :: collection
              (gen/fmap (fn [x] [coll x])
                        (gen/elements coll)))))
(gen/generate gen-collection-and-element 20)
[[-1 40 3 6199 -77 -433763 -412 46055 0 -6 0] 40]
```

2

3

4

5

6

7

8

10 11

```
(def gen-collection-and-element
2
      (gen/let [coll (gen/not-empty
3
                      (gen/vector gen/large-integer))
4
                     (gen/elements coll)]
        [coll x]))
    (gen/generate gen-collection-and-element 20)
    [[-37830 3546 -210996 3 0 -35206 6 517097] 3546]
```

5

6

Example: a 2d matrix

```
:: Bad, because the inner vectors
2
     ;; can have different lengths.
3
     (gen/vector (gen/vector gen/large-integer))
4
5
6
     ;; Generate a width first, then use bind to generate
7
     ;; a collection of vectors with the same width.
8
     (gen/let [width gen/nat
9
               rows (gen/vector
10
                     (gen/vector gen/large-integer
11
                                  width))]
12
      rows)
```

Summary: Generators, in General

- There are concrete data generators, and abstract combinators
- The abstractness is in service of the value proposition
- Like FP familiarity takes practice

Building

37 / 90

Long-Running Example: Dropbox

- Generate a random (nested) directory of files
- Generate changes to a directory

gen-file-name

2

3

4 5

gen-file-name

```
1  (gen/sample gen-file-name 1000)
2  ;; ExceptionInfo Couldn't satisfy such-that
3  ;; predicate after 10 tries.
4  ;; clojure.core/ex-info (core.clj:4725)
```

Common culprit in spec: (s/and s1 s2 s3 ...)

gen-file-name-2

2

3

4 5

gen-file-contents

1

2

3 4

5

6

8

9

10

1112

13

```
(def gen-file-contents
 gen/bytes)
(gen/sample gen-file-contents)
(#bytes
#bytes "Oe"
#bytes "8b"
#bytes "80e1d5"
#bytes "8ad6"
#bytes ""
#bytes "b0a0224119c8"
#bytes "f39a97ff"
#bytes "8a"
#bytes "22908190fc09eca901")
```

gen-file-metadata

2

3

4 5

gen-file-metadata

```
(def gen-datetime
 2
       (gen/fmap #(java.time.Instant/ofEpochMilli %)
 3
                 gen/large-integer))
 4
 5
     (gen/sample gen-datetime)
 6
     (#inst "1970-01-01T00:00:00.000Z"
 7
      #inst "1969-12-31T23:59:59.9997"
 8
      #inst "1969-12-31T23:59:59.9987"
 9
      #inst "1970-01-01T00:00:00.0007"
10
      #inst "1969-12-31T23:59:59.998Z"
11
      #inst "1970-01-01T00:00:00.007Z"
12
      #inst "1970-01-01T00:00:00.000Z"
13
      #inst "1969-12-31T23:59:59.938Z"
14
      #inst "1969-12-31T23:59:59.999Z"
15
      #inst "1970-01-01T00:00:00.035Z")
```

gen-file-metadata

1

2

3

4

5

6

7 8

10

11

```
(def gen-metadata
  (gen/hash-map :permissions gen-permissions-octal
                            gen/large-integer
               :user-id
               :group-id gen/large-integer
               :created-at gen-datetime
               :modified-at gen-datetime))
(gen/generate gen-metadata)
{:permissions "161",
 :user-id 484817,
 :group-id 10350453,
 :created-at #inst "1970-01-01T00:01:03.269Z",
 :modified-at #inst "1970-01-01T00:00:00.009Z"}
```

gen-directory

1

2

3

4

5

6 7

8

10 11

12

```
(defn gen-directory-of
  [gen-content]
  (gen/map gen-file-name
           (gen/hash-map :metadata gen-metadata
                          :content gen-content)))
(def gen-directory
  ;; use gen/such-that to filter out
  :: top-level byte arrays
  (gen/such-that map?
                 (gen/recursive-gen
                  gen-directory-of
                  gen-file-contents)))
```

gen-directory

12345678

9

10

11

12

13

14

15

16

17

18

19

20

```
(gen/generate gen-directory 10)
{"eqn" {:metadata {:created-at #inst "1970-01-01T00:00:00.003Z",
                   :group-id 0,
                   :modified-at #inst "1969-12-31T23:59:59.999Z"
                   :permissions "045",
                   :user-id 0}.
        :content {")" {:metadata {:created-at #inst "1969-12-31T23:59:
                                   :group-id 0,
                                   :modified-at #inst "1970-01-01T00:00
                                   :permissions "760",
                                   :user-id 0}.
                       :content #bytes "c5a1159d"}}},
 "xu_" {:metadata {:created-at #inst "1970-01-01T00:00:00Z",
                   :group-id 0,
                   :modified-at #inst "1969-12-31T23:59:59.999Z"
                   :permissions "372",
                   :user-id 0}.
        :content {"M" {:metadata {:created-at #inst "1969-12-31T23:59:
                                  :group-id -1,
```

gen-directory

1 | (def a-good-directory *1)

```
1 (def gen-directory-with-changes ????)
```

2

3

4 5

6

7

8

10

```
(defn gen-changes
  [directory]
 ????)
(def gen-directory-with-changes
  (gen/bind gen-directory
            (fn [directory]
              (gen/fmap (fn [changes]
                           {:directory directory
                            :changes changes})
                         (gen-changes directory)))))
```

1

2

3

4

5

6

7

8

10

11

12

```
(def gen-directory-with-changes
  (gen/bind gen-directory
            (fn [directory]
              (gen/fmap (fn [changes]
                          {:directory directory
                           :changes changes })
                        (gen-changes directory)))))
:: same as
(def gen-directory-with-changes
  (gen/let [directory gen-directory
            changes (gen-changes directory)]
   {:directory directory
     :changes changes}))
```

```
1  (defn gen-changes
2    [directory]
3    (gen/vector (gen-appension directory)))
```

```
(defn gen-appension
[directory]
;; 1. get all filepaths
;; 2. select (generate) a filepath
;; 3. generate random bytes to append
;; 4. package together and return
)
```

2

3

4

5

6

8

9 10

```
(defn all-filepaths
  [directory]
  (apply concat
         (for [[name {:keys [content]}] directory]
           (if (map? content)
             (map #(str name "/" %)
                  (all-filepaths content))
             [name]))))
(all-filepaths a-good-directory)
("eqn/)" "xu /M" "xu /X")
```

1

2

3

4

5

6

7

8

9

10 11

12 13

```
(defn gen-file-appension
  [directory]
  (let [filepaths (all-filepaths directory)]
    (assert (not-empty filepaths))
    (gen/fmap (fn [[filepath more-bytes]]
                {:filepath filepath
                 :append more-bytes})
              (gen/tuple (gen/elements filepaths)
                         gen/bytes))))
(gen/generate
 (gen-file-appension a-good-directory))
{:filepath "xu_/X",
 :append #bytes "3b5cfeb016458d"}
```

All together now...

2345678

9

10

11 12

13

14

15

16

17

18

```
(defn gen-changes
  [directory]
  ;; do something different if it's empty
  (gen/vector (gen-file-appension directory)))
(def gen-directory-with-changes
  (gen/let [directory gen-directory
            changes (gen-changes directory)]
    {:directory directory
     :changes changes}))
(gen/generate gen-directory-with-changes 3)
{:directory {"" {:metadata {:permissions "217",
                            :user-id -1,
                            :group-id 0,
                            :created-at #inst "1970-01-01T00:00:00.000
                            :modified-at #inst "1969-12-31T23:59:59.99
                 :content #bytes ""}},
 :changes [{:filepath "", :append #bytes ""}]}
```

Summary: Building Things

Just like FP, building generators takes practice

Fine-Tuning

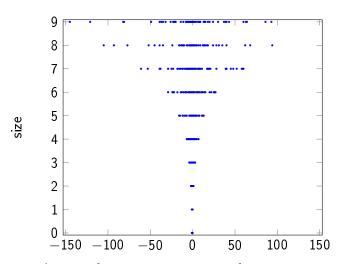
Fine-tune what, exactly?

Fine-tune what, exactly?

- Distribution/Growth
- Shrinking

Fine-Tuning Distribution/Growth

gen/large-integer



Distribution of gen/large-integer for size=0..9

What does size mean exactly?

Its meaning is determined by its use:

• Test runs use 0,1,2,...,199,0,1,2,...

size gotchas

- Testing with :num-tests < 200
- (gen/sample g) is size=0..9
- (gen/generate g) is size=30

gen/scale

"Big" is relative.

string vectors

```
1  (def g (gen/vector gen/string-ascii))
2
3  (repeatedly 5 #(gen/generate g 10))
4
5  (["V5N8'!)/" "_{^)X2ml"]
6  ["<h}>YmK^$2" "" "" "B-c>M/" ""]
7  ["b.e2I*$76"]
8  ["Vo" "Sk"]
9  ["u"])
```

```
(gen/generate g 200)
["s9&\"(5v?%{X6AuFU2dkM\"m0kb+1c(-mb@]C)n,$)g|/joLvU~W,g~L"
"&_<#bCsH7A-~\"_V=(r(i|G-'9bF.}N*NYass;*sWt(g06Y7XRZ37)6aQ:)lm~xgq(%xKIuZ>2V<+4dxsW7[eNygK!0b{\psi})R/o>()E
".{?T-@psGc%VDrQ{R;>5u0*/OdWfN=)\\;)3H}Ni>VOY!R63C@[^LdVZ?WOnAs |HemOvO("
"$~M(2ppm'H|*x$Z.0m39<2WbXbwbib= V2o^Jv-vT~~7BRci]$?vc8/1)"
"u~O>dYQG7>849<%DHT11iFUk>TcUMJ'hkk*\"M\\KX1.shq8ExD=KsQ7J(mI{i5Mh8<wG'n{|eL9e%c'K){,*PtYDW|pKV?,uouDsK
"GoP>[iISnY*wEGjBB^)+@(GW1$]h%aR.{JYMKq*X .]C6ty8e#%Y9.e:>GlE(RN;f2A\\5/om?o)%"
"h1=Z:6cJed'c\"% G+H.gg?@oS>"
"x[I]i1K+te6-}KTqQU<c1"
"%5$\\u(i, >uwWz'evt^Q?=m%H/t$~UB+_GpCdrTV?g3YJM2u$C,&1f.2Ym\\Q'18H>] a35FA4%Ih,\\e~Bv\\!Qfr6Kcl;'>T5'9D
"&_w5s69S%M:ZWU^q~hoYva%_$hDiraW~auq0>P#G-].z 55w&"
"'-OM=-uP];m8U?Q2(PsuPEp{F-NV'liq_\\,[K8.ABaTQI'MoqbU' _jGXz!$t$qm}!j~V!1'Y+GU6*q7b7v-y\\MR,*"
"i<d-b0_g(m~iMkLq4]@#srM9C(g&'J%bz.sA+BNX~OCj\"J&}&]).oMFA.%-tyBG]bCZ%h1D,]*~R~{2C,"
"pl]pVTI wVD4u]t)1,Et&)r9Tj,: 7+XzC8gRhWe8Ioc41102s?+Ueew>U{U(1x\"c<<1hSTs3ak.LbX3g}{;s16w_',C|1*s)rbA:
"2$XMP\\\"GE:xTkLBu}RX@LO S/S7[ziMI%GEOZY]I'drF<5oDTi5&EOgN:10Ne}}!c2^90<JR1<%T ArU1voHI\\$%:"
"LYmx,_u;,&n4%:>U-q>WC1]WGg)06}w"
```

"A~'Oh~?F@LG/}+-1\\ds1+X{Inofdj:z=+9S+/SuaLZM<rXf&u~Z>-=7YJ94Zzy[:.e;uDIsyDIsr4~iB(rxi<#E)Ck[H*C_tEIYW* ">r.= 3&E1Z.<T(hW\"2L%UG:ur=EmWA:60wE2zPbfW19|S(*!E=8K5snV!0>T='\$=mci&t&!9T.WK[z9\"'X~.'VHyW!=D!XhC~VF6 "{Z |np!C|F.b*H+t'?wgDefL/<=YHsRt<U=)vX{bk@NSR[UWJx&%J':uY4tu![]FIB8s6)9Pr@[*Y9h8st7z}\"pQ%7s,V]6{/bZM)M

"Zpw|d)R|{C\\e&XfN*|,K@|. N.6W?N'>p5N:@>UBLx4h,G5xRBPM^!Fz^zm]zMV[ufJg.\"6wLjBr{#'5t)1zkbbHk=!C\$&5B-m<g

"H\"EH\MpHD6UP,_@vO4kd9rDeK]Tk!H]M5=xSv3(e}_W3Y3OR\zkI'^GFgXaR.YIUOOUW'-CjVZQ.Duw3\$8=;tq#Tje.Y

"* B~NaUS/DhC7Z&A</:7cm\"A>&V7\$WX&=\\v&-{Vn=va?[>Z(%o N%C)ta7%/ .1U3dn{a8.SP\\8'/t"

"Z\"5.e#Cew{k5>^vm?YyUSDTL6;.^^g)Qx(J2I\$E5.g@!vM|->3 4\"cE)?LNW604Q>5P}m&A#."

69 / 90

"AU(%)Be>)WJZ4a)40:P?Z?=!Mh> Γ >>wf K7a&H"

string vectors

```
(def g-2 (gen/scale \#(/ \% 20)
2
                          (gen/vector gen/string-ascii)))
3
4
     (repeatedly 5 #(gen/generate g-2 200))
5
6
     (["Qw+b"]
7
     [".S." "C'NL.k2"]
8
      ["Sh?7?Wv" "" "@,~d_^*z" "z<" "]Q!" "XPVS"
10
      "/|@)WYvbH" "(=jTr" "ZQ}k[as;o0" "p&ri~;flQ6"]
      ["0=]{-=\" "BTqjqWg" ":Nlz" "F6I*(P"])
11
```

PSA



test.check / TCHECK-106

Variable-sized collection generators have exponential sizing issues when composed

Too Weird / Not Weird Enough

```
1  ;; doesn't often generate
2  ;; nontrivial text files
3  (def gen-file-contents
4  gen/bytes)
```

What's the best amount of weird?

```
;; doesn't often generate
2
    ;; nontrivial text files
3
    (def gen-file-contents
4
      gen/bytes)
5
6
    ;; half random bytes, half UTF-8
7
    (def gen-file-contents-2
8
      (gen/one-of [gen/bytes
9
                   gen-UTF8-bytes]))
```

```
;; Only generates small files
 2
     (def gen-file-contents-2
 3
       (gen/one-of [gen/bytes
4
                     gen-UTF8-bytes]))
 5
 6
     ;; Occasionally generates big files!
 7
     (def gen-file-contents-3
 8
       (gen/frequency [[45 gen/bytes]
                        [45 gen-UTF8-bytes]
10
                        [10 (gen/scale
                             (fn [size]
12
                                (if (<= 100 size)
                                  (* size size)
13
14
                                  size))
15
                             gen/bytes)]]))
```

9

11

gen-datetime

```
(def gen-datetime
 2
       (gen/fmap #(java.time.Instant/ofEpochMilli %)
 3
                 gen/large-integer))
 4
 5
     (gen/sample gen-datetime)
 6
     (#inst "1970-01-01T00:00:00.000Z"
 7
      #inst "1969-12-31T23:59:59.9997"
 8
      #inst "1969-12-31T23:59:59.9987"
 9
      #inst "1970-01-01T00:00:00.0007"
10
      #inst "1969-12-31T23:59:59.998Z"
11
      #inst "1970-01-01T00:00:00.007Z"
12
      #inst "1970-01-01T00:00:00.000Z"
13
      #inst "1969-12-31T23:59:59.938Z"
14
      #inst "1969-12-31T23:59:59.999Z"
15
      #inst "1970-01-01T00:00:00.035Z")
```

gen-datetime-2

```
(def gen-datetime-components
2
      (gen/hash-map
3
       :year (gen/fmap #(+ % 2017) gen/int)
4
       :month (gen/large-integer* {:min 1, :max 12})
5
       :day (gen/large-integer* {:min 1, :max 31})
6
       :hour (gen/large-integer* {:min 0, :max 23})
7
       :minute (gen/large-integer* {:min 0, :max 59})
8
       :second (gen/large-integer* {:min 0, :max 59})
       :millis (gen/large-integer* {:min 0, :max 1000})))
```

gen-datetime-2

1

2

3

4

5

6

7

8

9

10 11

12

13

14

```
(defn construct-datetime
  [{:keys [year month day
           hour minute second millis]}]
  (try
    (java.time.Instant/parse
     (format "%04d-%02d-%02dT%02d:%02d:%02d.%03dZ"
             year month day
             hour minute second millis))
    (catch Exception e
      ;; kind of dumb, but it works and it's easy
      (java.time.Instant/parse
       (format "%04d-%02d-%02dT%02d:%02d:%02d.%03dZ"
               year month 28
               hour minute second millis)))))
```

gen-datetime-2

```
(def gen-datetime-2
 2
       (gen/fmap construct-datetime
 3
                 gen-datetime-components))
 4
 5
     (gen/sample gen-datetime-2)
     (#inst "2017-02-02T01:01:01.001Z"
 6
 7
      #inst "2016-01-02T01:01:01.0007"
 8
      #inst "2017-02-01T00:01:00.0017"
 9
      #inst "2015-02-01T00:03:01.0037"
10
      #inst "2016-07-02T01:00:00.002Z"
11
      #inst "2022-03-02T01:02:03.001Z"
12
      #inst "2020-05-03T04:00:01.000Z"
13
      #inst "2019-02-09T10:05:11.000Z"
14
      #inst "2020-04-08T12:27:04.053Z"
15
      #inst "2013-08-06T21:04:03.001Z")
```

Fine-Tuning Shrinking

gen-datetime vs gen-datetime-2

gen/uuid

```
1  (def ^{:added "0.9.0"} uuid
2    "Generates a random type-4 UUID. Does not shrink."
3    (no-shrink
4    ...))
```

gen/bind

1

2

3

4

5

6

7 8

9

10 11

12

13 14

```
(def gen-matrix
  (gen/let [width gen/nat
            rows (gen/vector
                  (gen/vector gen/large-integer
                               width))]
    rows))
(quick-check 10000
             (prop/for-all [matrix gen-matrix]
               (->> matrix
                     (apply concat)
                     (not-any? #{42}))))
;; fails on a large matrix, shrinks to [[0 42 0]]
```

Custom Shrinking

Summary: Tweaking Things

- Things can be too big or small; gen/scale can help
- You can target specific sensitivities with gen/one-of and gen/frequency
- Modeling the domain better can help
- Shrinking is hard

Welp!

That's About It

- Generator combinators are abstract and declarative, to support growth and shrinking
- With practice, you can generate anything, and customize its distribution, growth, and shrinking

Thanks!