COMP 150PP: Language Design for the monad of measures

November 2, 2016

Prelude: algebra of characteristic functions

(At the board)

Measures (background)

A *measure* is a mathematical object that may be defined in two ways:

- A measure may be defined as a function from a *measurable set* to a nonnegative extended real number (zero to infinity). The function must be *countably additive*: the measure of a countable union of disjoint sets is the sum of their measures.

  The choice of measurable sets depends on the applications, but in a discrete space, typically every set is measurable, and on the real line, we typically take the *Borel sets*, which are formed by taking countable unions, intersections, and complements of intervals. (At higher dimensions, a slightly more complicated construction is required, but the ideas are the same.)

  In every case, the empty set is measurable and has measure zero, and the complement of the empty set—which is to say the entire space—is measurable. Sets with infinite measure are not merely permissible; they are commonplace.

- A measure may be defined as an *integrator*, which is a higher-order function. In the simple case, a measure of type `M a` defines an integrator that can be applied to any function f whose type is `a -> Real`.![The codomain ("range") of f may actually be any vector space over the real numbers, but if you understand the integration of real functions, it's The type variable `a` may be instantiated at a continuous space, a discrete space, or a hybrid.

  An integrator $[\![m]\!]_I$, when applied to function f, computes the integral of f according to the measure denoted by m. Like all integrals, $[\![m]\!]_I$ is a linear operator, and so it satisfies the following algebraic laws:

$$[\![m]\!]_I(f+g) = [\![m]\!]_I(f) + [\![m]\!]_I(g)$$
$$[\![m]\!]_I(c \cdot f) = c \cdot [\![m]\!]_I(f)$$

These two definitions are equivalent:

- Given an integrator, you can recover a measure function by integrating the characteristic function of a measurable set.
- Given a measure function, you can integrate any nonnegative measurable function by a countable sequence of approximations. The integrand f is approximated below by a sum of characteristic functions, where the domain of each function is a finite interval. To integrate an arbitrary function, integrate the positive and negative parts separately.

Lebesgue measure and counting measure

Lebesgue measure is the unique measure on the real line that maps each interval to its length. Integration against the Lebesgue measure is usually called "Lebesgue integration." Whenever the Lebesgue integral of a function is defined, the Riemann integral is also defined, and they are equal.

Given any countable set of points, the counting measure is the unique measure that assigns measure 1 to each point. The integrator takes a (countable) sum.

Probability distributions and "unnormalized distributions"

A probability distribution is a measure that assigns the entire space a measure of 1.

An "unnormalized distribution" is a measure that assigns the entire space a finite, nonzero measure.

Given any finite, nonzero measure, we can produce the expectation of a function f using the following equation:

$$E(f) = \frac{[\![m]\!]_I(f)}{[\![m]\!]_I(\lambda\_.1)}.$$

Language design and the monad of measures.

Today's plan is to think about language design and the monad of measures. In addition to the usual unit and bind operations, we have our familiar `u` operator (for the uniform distribution over the unit interval), plus the Lebesgue measure:

`u :: M Real`	Integrates a function over the unit interval
`Λ :: M Real`	Integrates a function over the real line (the Lebesgue measure)
`ℕ :: M Nat`	Counting measure on the natural numbers

We also have

```
observe :: Bool -> M ()
```

You might think about what mathematical object corresponds to a function of type `() -> Real`, and what it might mean to integrate such a function.

Today's exercise has several parts:

- Using what we know about the monad of probability distributions, define an "integrator semantics" for the monad of measures
- Using what we know about integration and about operations on integrators, extend our monadic language to be able to express such operations
- Figure out where and how to use probability-density functions

Integrator semantics

Using some of the techniques of denotational semantics, we'll define the meaning of a monadic language by translating each monadic term into an integrator:

1. $[\![\mathtt{u}]\!]_I(f) = \int_0^1 f(x)dx$
2. $[\![\Lambda]\!]_I(f) =$
3. $[\![\mathbb{N}]\!]_I(f) =$

   Here f has type `Nat -> Real`, where `Nat` is the type of natural numbers.
4. $[\![\mathtt{return}\ v]\!]_I(f) =$

   (Recall that if v is a value of type `a`, then `return` v has type `M a`, and therefore it can integrate any nonnegative measurable function of type `a -> Real`. When you have an idea, verify that it satisfies the algebraic laws for integrators.)
5. $[\![m >\!\!>\!\!= (\lambda x.m')]\!]_I(f) =$

   Here are the types:

m	`M a`
x	`a`
m'	`M b`
f	`b -> Real`

   An alternative way to write the term is `do` $\{x \leftarrow m; m'\}$.
6. $[\![\mathtt{observe}\ v]\!]_I(f) =$

   I want to alert you that in the presence of continuous variables this one is dodgy, but for today it shouldn't matter.

   Here are the types:

v	`Bool`
`observe` v	`M ()`
f	`() -> Real`

   It may help to know that the term

   ```
   observe e >>= \() -> m
   ```

   is more commonly written

   ```
   do { observe e; m }
   ```

Vector-space properties

The integral of a function~f can be an element of a vector space. In particular, this means that two integrals can be summed and that an integral can be multiplied by a real constant. Also, every vector space has a zero element.

7. Extend our language with a term whose denotation (translation) involves the sum of two integrals.
8. Extend our language with a term whose denotation involves multiplying an integral by a constant.
9. Extend our language with a term whose denotation involves the zero element of the vector space.

Application in practice

10. Write a term in the language to describe the distribution of rolls of a four-sided die. The term need not denote a probability measure.
11. Using your previous answer and your equations above, verify that the expected value of rolling a 4-sided die is 2.5.
12. Write a term in the language to describe a biased coin that comes up heads with probability 2/3 and tails with probability 1/3.

    ```
    data Coin = H | T
    ```
13. Using your previous answer and your equations above,

    - Write a term that models flipping the biased coin twice.
    - Using your equations, calculate the expected value of the function

      ```
      \(face, face') -> if face == face'
      then 0 else 1
      ```
    - Verify, by informal argument, that your answer gives the probability of getting different faces on two throws.

Probability-density functions

14. Consider a probabilistic computation in which a fair coin is flipped until it comes up heads.

    - Define a suitable probability-density function for use with measure $\mathbb{N}$.
    - Define an integrand function f whose expected value is the probability that at most three flips are required. Calculate $E(f)$.
    - Define an integrand function g whose expected value is the expected number of flips required to get heads.
    - Define a term in our language that can be used to integrate either f or g to get the right answers. In other words, figure out how to code the probability-density function.