Combinator Parsing Activity  
Mattox Beckman

<table>
<thead>
<tr>
<th>Manager</th>
<th>Keeps team on track</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recorder</td>
<td>Records decisions</td>
</tr>
<tr>
<td>Reporter</td>
<td>Reports to Class</td>
</tr>
</tbody>
</table>

Purpose

Monadic combinator parsers work very similarly to the LL parsers we covered before, but the monadic interface manages the input stream for us. The resulting parsers are much easier to read and to write. Your goals are:

- Understand the types of the parser combinators.
- Explain the result of executing a parser.
- Explain the `<|>` combinator.
- Implement many and many1.

Part 1 — The Types

```haskell
newtype Parser t = Parser (String -> [(t, String)]
run (Parser p) = p

oneOf xx = Parser (\inp -> case inp of
  (s:ss) | s `elem` xx -> [(s:ss)]
  otherwise         -> [])

sat pred = Parser (\inp -> case inp of
  (s:ss) | pred s    -> [(s:ss)]
  otherwise       -> [])

p1 = run (oneOf "abc") "axy"
p2 = run (oneOf "abc") "xya"
```

The newtype is like data, but the resulting type has only one constructor, and it is optimized away by the compiler. We use it instead of simply saying type Parser t = String -> [(t, String)] because we can’t declare a type as an instance, but we can declare a newtype as an instance.

Problem 1) What will be the values of p1 and p2?
Problem 2) Can you write the function `digit` that parses a digit? Use `sat` to do this. For more of a challenge, have it return an actual integer.

Part 2 — The Type Classes

```haskell
instance Functor Parser where
  fmap f (Parser p1) =
    Parser (inp -> [(f t, s) |
                    (t,s) <- p1 inp])

instance Applicative Parser where
  pure a = Parser (inp -> [(a,inp)])
  (Parser p1) <*> (Parser p2) =
    Parser (inp -> [(v1 v2, ss2) |
                    (v1,ss1) <- p1 inp,
                    (v2,ss2) <- p2 ss1])

instance Monad Parser where
  (Parser p) >>= f =
    Parser (inp -> concat [run (f v) inp' |
                           (v,inp') <- p inp])

data Exp = IntExp Integer
          | PlusExp Exp Exp
  deriving Show

p3 = run (IntExp <$> digit) "123"
p4 = run (PlusExp <$> getIntExp <*> getIntExp) "123"
p5 = do i1 <- getIntExp
       i2 <- getIntExp
       return (IntExp i1 i2)
```

Problem 3) What is the value of `p3`? Trace through the evaluation and be sure everyone on your team understands how we got that result.

Problem 4) Write the function `getIntExp` that is like `digit` but encapsulates the digit in an `IntExp`. 


Problem 5) What is the value of $p_4$? Trace through the evaluation and be sure everyone on your team understands how we got that result.

Problem 6) What is the value of $p_5$? Trace through the evaluation and be sure everyone on your team understands how we got that result.

Part 3 — Choice, Many, Many

1. $(\text{Parser } p_1) <|> (\text{Parser } p_2) =$
2. \[ \text{Parser } (\text{\inp} -> \text{take } 1 $ p_1 \text{ inp } ++ p_2 \text{ inp}) \]
3. \[ \text{string } [] = \text{Parser } (\text{\inp} -> [([[], \text{inp}])] \]
4. \[ \text{string } (s: ss) = \text{do } v <- \text{char } s \]
5. \[ \text{vv <- string } ss \]
6. \[ \text{return } $ v:vv \]
7. \[ \text{getPlusExp = do } \text{string } "+" \]
8. \[ e_1 <- \text{getExp} \]
9. \[ e_2 <- \text{getExp} \]
10. \[ \text{return } (\text{PlusExp } e_1 e_2) \]
11. \[ \text{getExp = getIntExp} \]
12. \[ <|> \text{getPlusExp} \]

Problem 7) Examine the code for $<|>$. How does it work? Hint: consider the cases that $p_1$ succeeds, $p_1$ fails but $p_2$ succeeds, and both $p_1$ and $p_2$ fail.
Problem 8) Write the parsers many p and many1 p that take zero or more (for many) or one or more (for many1) repetitions of p.

Problem 9) The way we suggested writing getIntExp only works for a single digit. Can you make it work for multi-digit integers now?

Part 4 — Precedence

Problem 10) Now add TimesExp to the mix. Stratify the grammar so that TimesExp has higher precedence than PlusExp.