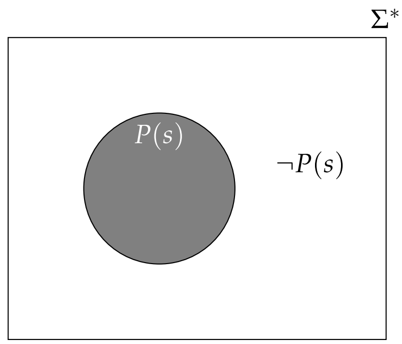
Complement Sets

Consider the following picture:



where Σ^* is in our case the set of all strings (what follows also holds for any kind of "domain", like the set of all integers or the set of all binary trees, etc). Let us assume $P(s)$ is a property that is about strings, for example $P(s)$ could be "the string s has an even length", or "the string s starts with the letter `a`". Every such property carves out a subset of strings from Σ^*, which in the picture above is depicted as a grey circle. This subset of strings is often written as a comprehension like

$$\{s \in \Sigma^* \mid P(s)\} \tag{1}$$

meaning all the s (out of Σ^*) for which the property $P(s)$ is true. If $P(s)$ would not be true then the corresponding string s would be outside the grey area where $\neg P(s)$ holds. Notice that sometimes the property $P(s)$ holds for every string in Σ^*. Then the grey area would fill out the whole rectangle and the set where $\neg P(s)$ holds is empty. Similarly, if the property $P(s)$ holds for no string, then the grey circle is empty.

Now, we are looking for the complement of the set defined in (1). This complement set is often written as

$$\overline{\{s \in \Sigma^* \mid P(s)\}}$$

It is the area of Σ^* which isn't grey, that is Σ^* minus $\{s \in \Sigma^* \mid P(s)\}$, **or** written differently it is the set $\{s \in \Sigma^* \mid \neg P(s)\}$. That means it is the set of all the strings where $\neg P(s)$ holds. Consequently we have for any complement set the equation:

$$\overline{\{s \in \Sigma^* \mid P(s)\}} = \{s \in \Sigma^* \mid \neg P(s)\} \tag{2}$$

Semantic Derivative

Our semantic derivative $Der\ c\ A$ is nothing else than a property that defines a subset of strings (inside Σ^*). The corresponding property $P(s)$ is $c :: s \in A$ because we defined $Der\ c\ A$ as

$$Der\ c\ A \stackrel{\text{def}}{=} \{s \in \Sigma^* \mid c :: s \in A\}$$

That means $Der\ c\ A$ is some grey area inside Σ^*. Obviously which subset, or which grey area, we are carving out from Σ^* depends on what we choose for c and A.

Let us see how this pans out in a concrete example. For this let Σ^* not be the set of all strings, but only the set of strings upto a length of 3 over the alphabet $\{a, b\}$. That means Σ^* (or the rectangle in the picture above) consists of the strings

$$\Sigma^* = \left\{ \begin{array}{l} [] \\ [a], [b] \\ [aa], [ab], [ba], [bb] \\ [aaa], [aab], [aba], [abb], [baa], [bab], [bba], [bbb] \end{array} \right\}$$

If we set A to $\{[aaa], [abb], [aa], [bb], []\}$, then $Der\ a\ A$ is the subset

$$Der\ a\ A = \{[aa], [bb], [a]\}$$

which is given by the definition of $Der\ a\ A \stackrel{\text{def}}{=} \{s \in \Sigma^* \mid a :: s \in A\}$. Now lets look at what the complement of this set looks like:

$$\overline{Der\ a\ A} = \left\{ \begin{array}{l} [] \\ [b] \\ [ab], [ba] \\ [aaa], [aab], [aba], [abb], [baa], [bab], [bba], [bbb] \end{array} \right\} \tag{3}$$

This can be calculated by "subtracting" $\{[aa], [bb], [a]\}$ from Σ^*. I let you check whether I did this correctly. According to the equation in (2) this should also be equal to

$$\overline{Der\ a\ A} = \{s \in \Sigma^* \mid \neg(a :: s \in A)\}$$

Let us test in turn every string in Σ^* and see whether $a :: s$ is in A which we set above to

$$\{[aaa], [abb], [aa], [bb], []\}$$

This gives rise to the following table where in the first column are all the strings of Σ^* and in the second whether $a :: s \in A$ holds. The third column is the negated version of the second, namely $\neg(a :: s \in A)$ which is the same as $a :: s \notin A$.

$s \in \Sigma^*$	is $a :: s \in A$?	$\neg(a :: s \in A) \Leftrightarrow a :: s \notin A$
$[]$	no	yes
$[a]$	yes	no
$[b]$	no	yes
$[aa]$	yes	no
$[ab]$	no	yes
$[ba]$	no	yes
$[bb]$	yes	no
$[aaa]$	no	yes
$[aab]$	no	yes
$[aba]$	no	yes
$[abb]$	no	yes
$[baa]$	no	yes
$[bab]$	no	yes
$[bba]$	no	yes
$[bbb]$	no	yes

Collecting all the yes in the third column gives you the set in (3). So it works out in this example. The idea is that this always works out. ;o)

BTW, notice that all three properties are the same

$$\neg(a :: s \in A) \quad \Leftrightarrow \quad a :: s \notin A \quad \Leftrightarrow \quad a :: s \in \overline{A}$$

This means we have

$$\overline{Der\ a\ A} = Der\ a\ \overline{A} \tag{4}$$

I let you check whether this makes sense.

Not-Regular Expression

With the equation in (4) we can also very quickly verify that the *der*-definition for the not-regular expression satisfies the property of derivatives, namely

$$\forall c\, r. \quad L(der\ c\ r) = Der\ c\ (L(r)) \tag{5}$$

holds. We defined the language of a not-regular expression as

$$L(\sim r) \stackrel{\text{def}}{=} \Sigma^* - L(r)$$

Using the overline notation, maybe I should have defined this equivalently as

$$L(\sim r) \stackrel{\text{def}}{=} \overline{L(r)}$$

meaning all the strings that r *cannot* match. We defined the derivative for the not-regular expression as

$$der\ c\ (\sim r) \stackrel{\text{def}}{=} \sim (der\ c\ r)$$

The big question is now does this definition satisfy the property in (5)? Lets see: We would have to prove this by induction on regular expressions. When we are in the case for $\sim r$ the reasoning is as follows:

$$\begin{array}{llll}
L(der\ c\ (\sim r)) & \stackrel{\text{def}}{=} & L(\sim (der\ c\ r)) & \text{by definition of } der \\
 & \stackrel{\text{def}}{=} & \overline{L(der\ c\ r)} & \text{by definition of } L \\
 & = & \overline{Der\ c\ (L(r))} & \text{by IH} \\
 & = & Der\ c\ \overline{(L(r))} & \text{by (4)} \\
 & \stackrel{\text{def}}{=} & Der\ c\ (L(\sim r)) & \text{by definition of } L
\end{array}$$

That means we have established the property of derivatives in the $\sim r$-case…yippee ;o)