The Chain Rule and Friends in ACL2(r)

Ruben Gamboa
Department of Computer Science
University of Wyoming
Laramie, Wyoming
ruben@uwyo.edu

John R. Cowles
Department of Computer Science
University of Wyoming
Laramie, Wyoming
cowles@cs.uwyo.edu

Extended Abstract

ACL2(r) includes a theory of differentiation, and a collection of facts such as $(x^n)' = n \cdot x^{n-1}$. ACL2(r) also includes books with theorems, such as Taylor's Theorem, which make extensive use of differentiation. This paper reports on a new set of ACL2(r) books that justify the familiar algebraic differentiation rules:

- $(f+g)'(x) = f'(x) + g'(x)$.
- $(-f)'(x) = -f'(x)$.
- $(f \cdot g)'(x) = f(x)g'(x) + f'(x)g(x)$.
- $(1/f)'(x) = -f'(x)/(f(x) \cdot f(x))$.
- $(f \circ g)'(x) = f'(g(x))g'(x)$.
- $(f^{-1})'(x) = 1/f'(f^{-1}(x))$, where f^{-1} is the (compositional) inverse of f.

Recall that ACL2(r) uses nonstandard analysis to introduce the irrationals. The definition of differentiability in NSA is as follows.

Definition 1. *A function f is* differentiable *at a standard point x_0, if at all points x_1 and x_2 that are i-close but not equal to x_0, the following conditions hold:*

- $\frac{f(x_0)-f(x_1)}{x_0-x_1}$ *is i-limited.*
- $\frac{f(x_0)-f(x_1)}{x_0-x_1} \approx \frac{f(x_0)-f(x_2)}{x_0-x_2}$.

If f is differentiable *at the point x_0, we define its* derivative *as* $\left(\frac{f(x_0)-f(x_1)}{x_0-x_1}\right)^*$.

The notation y^* denotes the standard part of y, and $x \approx y$ denotes that x is i-close to y. The derivative is defined as the standard part of an arbitrary differential. The second differentiability condition ensures that the choice of differential does not matter, since all the differentials are i-close to each other, hence they have the same standard part.

Naturally, this definition is introduced in ACL2(r) using encapsulate. Once it is introduced, we can define the differential of f as $\Delta f(x,\epsilon) \equiv \frac{f(x+\epsilon)-f(x)}{\epsilon}$, and the derivative *df* as the unique standard function such that $df(x) = (\Delta f(x,\epsilon))^*$, where ϵ in the definition of *df* is chosen so that $x+\epsilon$ is in the domain of f. Then it is easy to show that $df(x) \approx \Delta f(x,\epsilon)$ whenever $\epsilon \approx 0$.

Suppose f_1 and f_2 are introduced as constrained functions that satisfy the differentiability criterion described above. The function $s(x) \equiv f_1(x) + f_2(x)$ can be easily defined, and it is trivial to show that it, too, is differentiable; i.e., that it satisfies the constraints of a differentiable function. The same goes for $u(x) \equiv -f_1(x)$ and $p(x) \equiv f_1(x) \cdot f_2(x)$.

Things are little trickier for $q(x) \equiv 1/f_1(x)$. The reason is that q is not differentiable at the point x when $f_1(x) = 0$. While it would be natural to add the hypothesis $f_1(x) \neq 0$ to the theorem that establishes the differentiability of $q(x)$, this would not be a good solution for ACL2, because this extra hypothesis would prevent $q(x)$ from meeting the constraints of a differentiable function, so it could not be used in a functional-instantiate. Instead, we introduce a new function $f_3(x)$ that is differentiable *and* guaranteed to be non-zero.

The Chain Rule states that $(f_1 \circ f_2)'(x) = f_1'(f_2(x))f_2'(x)$. The proof of this result is straight-forward, although care needs to be taken when $f_2'(x) = 0$. Notably, the proof uses the fact that f_2 is continuous, since it is differentiable, and this fact is easily established via functional instantiation.

Finally, when f_1 is an invertible function (i.e., 1-1 and onto), we can introduce the inverse function $g(y) \equiv f_1^{-1}(y)$ implicitly using the macro definv (described in another workshop paper). Again, $g(y)$ is differentiable, and its derivative is given by $g'(y) = 1/f_1'(f_1^{-1}(y))$. As was the case with quotients, this requires that $f_1'(x) \neq 0$ over the domain of f_1. This criterion was added to the constraints of f_1 (or more precisely, to a new function f_4), so that g satisfied the differentiability constraints with no extra hypotheses.

The reason we are concerned with extra hypotheses, of course, is that we want to use these theorems repeatedly to find derivatives of functions such as $\sin(x^2) - \cos(\ln x)$. This can be done with the theorems we've proved so far, but it is a tedious process. In the future, we would like to automate this process by writing a macro that defines the derivative of a given function and constructs the appropriate proof.

Another future direction is support for derivatives over $\mathbb{C}$. We are in the process of extending the current formalization to support complex curves, and we plan to formalize the differentiation of complex functions of complex variables in the future.

.