Formal Verification of Medina's Sequence of Polynomials for Approximating Arctangent

Ruben Gamboa and John Cowles

Department of Computer Science
University of Wyoming
Laramie, Wyoming 82071

{ruben,cowles}@uwyo.edu

ACL2 Workshop 2014
Vienna, Austria

Outline

Motivation

- ACL2 is great for formalizing various aspects of microprocessors
- Significant efforts have targeted floating-point instructions, e.g., FADD, FMUL, etc.
- ACL2(r) comes into play when reasoning about transcendental or irrational functions, e.g., SIN, SQRT
- The goal (driven by David Russinoff) is to reason about the x87 trigonometric, logarithmic, and exponential functions: FSIN, FCOS, FSINCOS, FPTAN, **FPATAN**, F2XM1, FYL2X, and FYL2XP1

Motivation

- To reason about a function such as FSIN, we require theorems that compare $\mathrm{FSIN}(x)$ and the true value $\sin(x)$
- At the very least, we need to compare $\mathrm{FSIN}(x)$ and some function $f(x)$ such that $|f(x) - \sin(x)| < \mu$, where μ is sufficiently small, as compared to the machine word size
- An attractive option is to use approximating polynomials as $f(x)$, e.g., Taylor polynomials
- This approach works for sine and cosine, but it is not very practical for arctangent
- The reason is that the Taylor series for arctangent converges very slowly

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}$$

Medina's Approximation

- Medina suggested a new polynomial approximation to arctangent
- It starts with the following sequence of polynomials

$$p_1(x) = 4 - 4x^2 + 5x^4 - 4x^5 + x^6$$

$$p_m(x) = x^4(1-x)^4 p_{m-1}(x) + (-4)^{m-1} p_1(x)$$

Medina's Approximation

- Medina suggested a new polynomial approximation to arctangent
- It starts with the following sequence of polynomials

$$p_1(x) = 4 - 4x^2 + 5x^4 - 4x^5 + x^6$$
$$p_m(x) = x^4(1-x)^4 p_{m-1}(x) + (-4)^{m-1} p_1(x)$$

- Using calculus, it can be shown that

$$\left| \int_0^x \frac{p_m(t)}{(-1)^{m+1} 4^m} dt - \arctan(x) \right| \leq \left| \frac{1}{4} \right|^{5m}$$

Medina's Approximation

- Medina suggested a new polynomial approximation to arctangent
- It starts with the following sequence of polynomials

$$p_1(x) = 4 - 4x^2 + 5x^4 - 4x^5 + x^6$$
$$p_m(x) = x^4(1-x)^4 p_{m-1}(x) + (-4)^{m-1} p_1(x)$$

- Using calculus, it can be shown that

$$\left| \int_0^x \frac{p_m(t)}{(-1)^{m+1} 4^m} dt - \arctan(x) \right| \leq \left| \frac{1}{4} \right|^{5m}$$

- That shows that the polynomial sequence $h_m(x)$ defined below converges quickly to arctangent on $[0, 1]$

$$h_m(x) \equiv \int_0^x \frac{p_m(t)}{(-1)^{m+1} 4^m} dt.$$

Another Motivation

- ACL2(r) knows a lot of facts about trigonometry and calculus, but there are many gaps
- Over the years, ACL2(r) those foundational gaps have narrowed
- The challenge for us is to determine whether the gaps are small enough to support verification
- I.e., we are interested in whether the bulk of the effort is related to reasoning about arctangent, or to prove more foundational lemmas

Outline

Introducing Arctangent to ACL2(r)

- Stock ACL2(r) defines the trigonometric functions and proves some common identities
- The tangent function is defined (using a macro) as

$$\tan(x) \equiv \frac{\sin(x)}{\cos(x)}$$

Introducing Arctangent to ACL2(r)

- Stock ACL2(r) defines the trigonometric functions and proves some common identities
- The tangent function is defined (using a macro) as

$$\tan(x) \equiv \frac{\sin(x)}{\cos(x)}$$

- Unfortunately, arctangent is not defined in the base ACL2(r)
- So the first task was to admit a definition using the principle of inverse functions

Inverse Functions in ACL2(r)

- In principle, the inverse function f^{-1} can be defined whenever f is 1-to-1
- In practice, ACL2(r) restricts f to being a strictly increasing (or decreasing) function on its range
- Worse, f must be continuous, and the user is required to provide lower and upper bounds for any $f^{-1}(x)$

Inverse Functions in ACL2(r)

- In principle, the inverse function f^{-1} can be defined whenever f is 1-to-1
- In practice, ACL2(r) restricts f to being a strictly increasing (or decreasing) function on its range
- Worse, f must be continuous, and the user is required to provide lower and upper bounds for any $f^{-1}(x)$

- As it turns out, inverse functions are defined constructively using a macro
- Essentially, ACL2(r) uses a bisection search using the lower and upper bound given by the user

Introducing Arctangent: Tangent is 1-to-1

- Following convention, we restrict the domain of tangent to $(-\pi/2, \pi/2)$
- We still need to show that tangent is 1-to-1 on this domain
- Our approach was to show that tangent is **increasing** on this domain
- That follows from the sign of the derivative of tangent
- And the derivative follows from the quotient rule
  - provided $\cos(x) \neq 0$ for $x \in (-\pi/2, \pi/2)$

Introducing Arctangent: Tangent is 1-to-1

- Following convention, we restrict the domain of tangent to $(-\pi/2, \pi/2)$
- We still need to show that tangent is 1-to-1 on this domain
- Our approach was to show that tangent is **increasing** on this domain
- That follows from the sign of the derivative of tangent
- And the derivative follows from the quotient rule
  - provided $\cos(x) \neq 0$ for $x \in (-\pi/2, \pi/2)$
  - which was previously proved (way back in 1998)

Introducing Arctangent: Tangent is 1-to-1

- Following convention, we restrict the domain of tangent to $(-\pi/2, \pi/2)$
- We still need to show that tangent is 1-to-1 on this domain
- Our approach was to show that tangent is **increasing** on this domain
- That follows from the sign of the derivative of tangent
- And the derivative follows from the quotient rule
  - provided $\cos(x) \neq 0$ for $x \in (-\pi/2, \pi/2)$
  - which was previously proved (way back in 1998)
- Bonus: differentiable functions are continuous, so tangent is continuous

Introducing Arctangent: Tangent is 1-to-1

- We discovered that ACL2 was missing an obvious foundational lemma:

Theorem

If $f'(x) > 0$ for all x in some interval I, then f is increasing on I.

- We had to prove this foundational lemma as part of this verification effort
- Our proof is based on the MVT (which seems a bit heavy-handed)

Introducing Arctangent: Lower and Upper Bounds

- Given an arbitrary y, we need x_1 and x_2 such that $\tan(x_1) \leq y \leq \tan(x_2)$

Introducing Arctangent: Lower and Upper Bounds

- Given an arbitrary y, we need x_1 and x_2 such that $\tan(x_1) \leq y \leq \tan(x_2)$
- Case 1: $0 \leq y \leq 1$.
  - Since $\tan(0) = 0$ and $\tan(\pi/4) = 1$, we have found $x_1 = 0$ and $x_2 = \pi/4$ that work

Introducing Arctangent: Lower and Upper Bounds

- Given an arbitrary y, we need x_1 and x_2 such that $\tan(x_1) \leq y \leq \tan(x_2)$
- Case 1: $0 \leq y \leq 1$.
  - Since $\tan(0) = 0$ and $\tan(\pi/4) = 1$, we have found $x_1 = 0$ and $x_2 = \pi/4$ that work
- Case 2: $y > 1$.
  - Observe that whenever $x \in [\pi/4, \pi/2)$, $\tan(x) = \sin(x)/\cos(x) \geq 1/(\sqrt{2}\cos(x))$
  - Let $x = \arccos(1/\sqrt{2}y)$
  - Then $\tan(x) \geq 1/(\sqrt{2}\cos(x)) = 1/(\sqrt{2}\cos(\arccos(1/(\sqrt{2}y)))) = 1/(\sqrt{2}(1/\sqrt{2}y)) = y$
  - So the choice $x_1 = 1$ and $x_2 = \arccos(1/\sqrt{2}y)$ yields $\tan(x_1) \leq y \leq \tan(x_2)$

Introducing Arctangent: Lower and Upper Bounds

- Given an arbitrary y, we need x_1 and x_2 such that $\tan(x_1) \leq y \leq \tan(x_2)$
- Case 1: $0 \leq y \leq 1$.
  - Since $\tan(0) = 0$ and $\tan(\pi/4) = 1$, we have found $x_1 = 0$ and $x_2 = \pi/4$ that work
- Case 2: $y > 1$.
  - Observe that whenever $x \in [\pi/4, \pi/2)$, $\tan(x) = \sin(x)/\cos(x) \geq 1/(\sqrt{2}\cos(x))$
  - Let $x = \arccos(1/\sqrt{2}y)$
  - Then $\tan(x) \geq 1/(\sqrt{2}\cos(x)) = 1/(\sqrt{2}\cos(\arccos(1/(\sqrt{2}y)))) = 1/(\sqrt{2}(1/\sqrt{2}y)) = y$
  - So the choice $x_1 = 1$ and $x_2 = \arccos(1/\sqrt{2}y)$ yields $\tan(x_1) \leq y \leq \tan(x_2)$
- Case 3: $y < 0$
  - Use $\tan(-x) = -\tan(x)$ and try Case 1 or Case 2 with $-x$ instead

Outline

Derivative of Arctangent

- The derivative of an inverse function is given by

$$\frac{df^{-1}(y)}{dy} = \frac{1}{\frac{df(f^{-1}(y))}{dx}}$$

- This only holds when f' is not infinitesimally small (including zero)

Derivative of Arctangent

- The derivative of an inverse function is given by

$$\frac{df^{-1}(y)}{dy} = \frac{1}{\frac{df(f^{-1}(y))}{dx}}$$

- This only holds when f' is not infinitesimally small (including zero)
- Can $\tan'(x)$ be infinitesimal?

Derivative of Arctangent

- The derivative of an inverse function is given by

$$\frac{df^{-1}(y)}{dy} = \frac{1}{\frac{df(f^{-1}(y))}{dx}}$$

- This only holds when f' is not infinitesimally small (including zero)

- Can $\tan'(x)$ be infinitesimal?

- We know $\tan'(x) = \sec^2(x) = 1/\cos^2(x) \geq 1/1 = 1$
- So $\tan'(x)$ is never infinitesimal, and we can use the formula above to find $\arctan'(x)$

Derivative of Arctangent

- Using the formula, we get

$$\arctan'(x) = \frac{1}{\sec^2(\arctan(x))}$$

- The trigonometric identity $\sec^2(x) = 1 + \tan^2(x)$ reduces this to

$$\arctan'(x) = \frac{1}{1 + \tan^2(\arctan(x))} = \frac{1}{1 + x^2}$$

Useful Lemma

- We know that

$$\arctan'(x) = \frac{1}{1 + \tan^2(\arctan(x))} = \frac{1}{1 + x^2}$$

Useful Lemma

- We know that

$$\arctan'(x) = \frac{1}{1 + \tan^2(\arctan(x))} = \frac{1}{1 + x^2}$$

- The Fundamental Theorem of Calculus allows us to turn this result around into

$$\int_a^b \frac{dx}{1 + x^2} = \arctan(b) - \arctan(a)$$

- We will use this result later today!

Useful Lemma

- We know that

$$\arctan'(x) = \frac{1}{1 + \tan^2(\arctan(x))} = \frac{1}{1 + x^2}$$

- The Fundamental Theorem of Calculus allows us to turn this result around into

$$\int_a^b \frac{dx}{1 + x^2} = \arctan(b) - \arctan(a)$$

- We will use this result later today!

- Note: Matt proved a version of the FTC back in 1999. As part of this project, we updated that result to make it more accessible and also consistent with the current definition of continuity and derivative

Outline

Calculus of Monomials ax^n

- ACL2 defines the function `expt` corresponding to x^n
- We could use induction to prove that the derivative of x^n is nx^{n-1}

Calculus of Monomials ax^n

- ACL2 defines the function `expt` corresponding to x^n
- We could use induction to prove that the derivative of x^n is nx^{n-1}

- But there are some significant challenges

- Approach #1: Use the nonstandard definition of derivative
  - We can only induct up to standard values of n, so we get a weaker result
  - Moreover, the key lemma is the product rule
  - And the product rule has a non-classical condition, so we run into problems with the free variable n

Calculus of Monomials ax^n

- ACL2 defines the function `expt` corresponding to x^n
- We could use induction to prove that the derivative of x^n is nx^{n-1}

- But there are some significant challenges

- Approach #1: Use the nonstandard definition of derivative
  - We can only induct up to standard values of n, so we get a weaker result
  - Moreover, the key lemma is the product rule
  - And the product rule has a non-classical condition, so we run into problems with the free variable n

- Approach #2: Use the classical definition of derivative
  - Now the hypotheses and conclusion of the theorem are implicit as constraints
  - And the constraints do not get automatically propagated during induction!

Calculus of Monomials ax^n: A Workaround

- ACL2 defines the function `expt` corresponding to x^n
- ACL2(r) defines the function `raise` as $e^{n \ln(x)}$, which is also equal to x^n

Calculus of Monomials ax^n: A Workaround

- ACL2 defines the function `expt` corresponding to x^n
- ACL2(r) defines the function `raise` as $e^{n \ln(x)}$, which is also equal to x^n

- ACL2(r) already knows the derivative of e^x and $\ln(x)$
- So getting the derivative should be a small matter of plumbing

Calculus of Monomials ax^n: A Workaround

- ACL2 defines the function `expt` corresponding to x^n
- ACL2(r) defines the function `raise` as $e^{n \ln(x)}$, which is also equal to x^n

- ACL2(r) already knows the derivative of e^x and $\ln(x)$
- So getting the derivative should be a small matter of plumbing

- Once the derivative of `raise` is known, the derivative of `expt` follows
- And the derivative of ax^n follows trivially from the product rule

Calculus of Polynomials

- ACL2(r) knows a little about polynomials represented as lists of coefficients with a special evaluator function, e.g., `(eval-polynomial poly x)`
  - Prior work on transcendental numbers

Calculus of Polynomials

- ACL2(r) knows a little about polynomials represented as lists of coefficients with a special evaluator function, e.g., `(eval-polynomial poly x)`
  - Prior work on transcendental numbers
- An advantage of using lists to represent polynomials is that it permits a sort of higher-order reasoning
- E.g., we can introduce functions that compute the derivative and integral of a polynomial

Calculus of Polynomials

- ACL2(r) knows a little about polynomials represented as lists of coefficients with a special evaluator function, e.g., `(eval-polynomial poly x)`
  - Prior work on transcendental numbers
- An advantage of using lists to represent polynomials is that it permits a sort of higher-order reasoning
- E.g., we can introduce functions that compute the derivative and integral of a polynomial
- Of course, we have to show that those (trivial) functions actually compute the corresponding derivative and integral of the evaluator function!

Calculus of Polynomials

- ACL2(r) knows a little about polynomials represented as lists of coefficients with a special evaluator function, e.g., `(eval-polynomial poly x)`
  - Prior work on transcendental numbers
- An advantage of using lists to represent polynomials is that it permits a sort of higher-order reasoning
- E.g., we can introduce functions that compute the derivative and integral of a polynomial
- Of course, we have to show that those (trivial) functions actually compute the corresponding derivative and integral of the evaluator function!
- Note: We are forced to use the classical ϵ-δ style definitions of derivative and integral, since the evaluator function has the free variable `poly`

Outline

A Key Polynomial Sequence

- Important Restriction: From now on, we assume that we're only interested in $\arctan(x)$ for values of $x \in [0, 1]$
- It is easy to extend this definition to all values of x, but the key results below only work for x in this range

A Key Polynomial Sequence

- Important Restriction: From now on, we assume that we're only interested in $\arctan(x)$ for values of $x \in [0, 1]$
- It is easy to extend this definition to all values of x, but the key results below only work for x in this range

- Medina starts by defining the following sequence of polynomials

$$p_1(x) = 4 - 4x^2 + 5x^4 - 4x^5 + x^6$$
$$p_m(x) = x^4(1 - x)^4 p_{m-1}(x) + (-4)^{m-1} p_1(x)$$

- We defined ACL2 functions of the p_m both explicitly and using `eval-polynomial`

Rewriting the Key Polynomial Sequence

- Medina provides an alternative definition of the p_m, which holds for $m \geq 2$:

$$p_m(x) = \frac{x^{4m}(1-x)^{4m} + (-4)^m}{1+x^2}$$

Rewriting the Key Polynomial Sequence

- Medina provides an alternative definition of the p_m, which holds for $m \geq 2$:

$$p_m(x) = \frac{x^{4m}(1-x)^{4m} + (-4)^m}{1+x^2}$$

- Notice that this alternative equation obscures the fact that p_m is a polynomial, so we do only use explicit definitions in ACL2

Rewriting the Key Polynomial Sequence

- Consider the alternative definition of p_m:

$$p_m(x) = \frac{x^{4m}(1-x)^{4m} - (-4)^m}{1+x^2}$$

- We focus now on the term $x(1-x) = x - x^2$
- How big can this term be?

Rewriting the Key Polynomial Sequence

- Consider the alternative definition of p_m:

$$p_m(x) = \frac{x^{4m}(1-x)^{4m} - (-4)^m}{1+x^2}$$

- We focus now on the term $x(1-x) = x - x^2$
- How big can this term be?

- We can answer that by using the Calc I algorithm:
  - take the derivative, set it to zero, and check the endpoints
- We find that $x(1-x) \leq 1/4$, so $x^{4m}(1-x)^{4m} \leq 1/4^{4m}$

Rewriting the Key Polynomial Sequence

- Consider the alternative definition of p_m:

$$p_m(x) = \frac{x^{4m}(1-x)^{4m} - (-4)^m}{1+x^2}$$

- We focus now on the term $x(1-x) = x - x^2$
- How big can this term be?

- We can answer that by using the Calc I algorithm:
  - take the derivative, set it to zero, and check the endpoints
- We find that $x(1-x) \leq 1/4$, so $x^{4m}(1-x)^{4m} \leq 1/4^{4m}$

- We did need to add a foundational result here, namely the First Derivative Test (which we also proved with the MVT)

Bounding More Terms

- Recall that $x^{4m}(1-x)^{4m} \leq 1/4^{4m}$
- Since $1 + x^2 \geq 1$, this immediately yields that

$$\frac{x^{4m}(1-x)^{4m}}{1+x^2} \leq \left(\frac{1}{4}\right)^{4m}$$

Bounding More Terms

- Recall that $x^{4m}(1-x)^{4m} \leq 1/4^{4m}$
- Since $1 + x^2 \geq 1$, this immediately yields that

$$\frac{x^{4m}(1-x)^{4m}}{1+x^2} \leq \left(\frac{1}{4}\right)^{4m}$$

- Integrating both sides results in an important bound:

$$\begin{aligned}\int_0^x \frac{t^{4m}(1-t)^{4m}}{1+t^2}dt &\leq \int_0^x \left(\frac{1}{4}\right)^{4m} dt \\ &= \left(\frac{1}{4}\right)^{4m} x \\ &\leq \left(\frac{1}{4}\right)^{4m}\end{aligned}$$

Bounding More Terms

- Aside: The last step required more foundational lemmas, namely the monotonicity of integrals

Putting It Together

- So far we have shown that

$$\begin{aligned}\frac{x^{4m}(1-x)^{4m}}{1+x^2} &= p_m(x) + \frac{(-4)^m}{1+x^2} \\ &= p_m(x) - \frac{(-1)^{m+1}4^m}{1+x^2}.\end{aligned}$$

Putting It Together

- So far we have shown that

$$\begin{aligned}\frac{x^{4m}(1-x)^{4m}}{1+x^2} &= p_m(x) + \frac{(-4)^m}{1+x^2} \\ &= p_m(x) - \frac{(-1)^{m+1}4^m}{1+x^2}.\end{aligned}$$

- Integrating both sides of that equation and dividing by the constant $(-1)^{m+1}4^m$ gives the following:

$$\left| \int_0^x \frac{p_m(t)}{(-1)^{m+1}4^m} - \frac{1}{1+t^2} dt \right| \leq \left| \frac{1}{4} \right|^{5m}$$

Putting It Together

- The final result follows from integrating the left-hand side

$$\left| \int_0^x \frac{p_m(t)}{(-1)^{m+1}4^m} - \frac{1}{1+t^2} dt \right| \leq \left| \frac{1}{4} \right|^{5m}$$

$$\left| \int_0^x \frac{p_m(t)}{(-1)^{m+1}4^m} dt - \arctan(x) \right| \leq \left| \frac{1}{4} \right|^{5m}$$

Putting It Together

- The final result follows from integrating the left-hand side

$$\left| \int_0^x \frac{p_m(t)}{(-1)^{m+1}4^m} - \frac{1}{1+t^2} dt \right| \leq \left| \frac{1}{4} \right|^{5m}$$

$$\left| \int_0^x \frac{p_m(t)}{(-1)^{m+1}4^m} dt - \arctan(x) \right| \leq \left| \frac{1}{4} \right|^{5m}$$

- Notice that the remaining integral is of a polynomial
- We can now define a new polynomial sequence:

$$h_m(x) \equiv \int_0^x \frac{p_m(t)}{(-1)^{m+1}4^m} dt$$

Putting It Together

- The final result follows from integrating the left-hand side

$$\left| \int_0^x \frac{p_m(t)}{(-1)^{m+1}4^m} - \frac{1}{1+t^2} dt \right| \leq \left| \frac{1}{4} \right|^{5m}$$

$$\left| \int_0^x \frac{p_m(t)}{(-1)^{m+1}4^m} dt - \arctan(x) \right| \leq \left| \frac{1}{4} \right|^{5m}$$

- Notice that the remaining integral is of a polynomial
- We can now define a new polynomial sequence:

$$h_m(x) \equiv \int_0^x \frac{p_m(t)}{(-1)^{m+1}4^m} dt$$

- The result is the fast polynomial approximation to arctangent:

$$|h_m(x) - \arctan(x)| \leq \left| \frac{1}{4} \right|^{5m}$$

Outline

Conclusions

- Medina's result has important applications for the verification of hardware implementations of arctangent
- We plan to pursue this effort in the coming year

- Medina's result uses a significant amount of calculus
- It turned out to be a great case study for ACL2(r)
- Although some foundational results were required, ACL2(r) appears to be getting stable enough for actual applications

Conclusions

- An interesting fact is that it's possible to prove Medina's series and the Taylor series converge to the same function
- Such an effort can be envisioned in ACL2 (and it may be part of a hardware verification project)
- But we do not believe that proof can really be carried out in ACL2
- The recommended approach (per the ACL2 help list) is to show that each of the two series give the same result
- But the function they converge to does not actually exist in ACL2!