Theory and Practice of Symbolic Integration in Maple

John May
Senior Developer

Mathematical Software

Maplesoft

September 2010

Introduction

When the talking version of a popular doll said “Math is Hard!”, she was almost certainly talking about symbolic integration.

Introduction

When the talking version of a popular doll said “Math is Hard!”, she was almost certainly talking about symbolic integration.

But! The integrals in my first year calculus class were all easy! Was it all lies?

Introduction

When the talking version of a popular doll said "Math is Hard!", she was almost certainly talking about symbolic integration.

But! The integrals in my first year calculus class were all easy! Was it all lies?

Yes.

Two Related Problems

Indefinite Integration (anti-differentiation)

$$F(x) = \int f(x)\,dx \iff \frac{d}{dx}F(x) = f(x)$$

Two Related Problems

Indefinite Integration (anti-differentiation)

$$F(x) = \int f(x)\,dx \iff \frac{d}{dx}F(x) = f(x)$$

Definite Integration (the area under a curve)

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=0}^{n} f(x_i)\,\Delta x$$

Two Related Problems

Indefinite Integration (anti-differentiation)

$$F(x) = \int f(x)\,dx \iff \frac{d}{dx}F(x) = f(x)$$

Definite Integration (the area under a curve)

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=0}^{n} f(x_i)\,\Delta x$$

Unfortunately, not always as related as we might hope.

Some History: SAINT and SIN

Some of the first attempts at symbolic integration software were done as Artificial Intelligence Research.

Some History: SAINT and SIN

Some of the first attempts at symbolic integration software were done as Artificial Intelligence Research.
SAINT (1961) by J. Slagle and SIN (1967) by J. Moses were both expert systems designed to compute definite integrals.
SAINT was purely heuristic search based and aimed to be as good as a first year student.

Some History: SAINT and SIN

Some of the first attempts at symbolic integration software were done as Artificial Intelligence Research.
SAINT (1961) by J. Slagle and SIN (1967) by J. Moses were both expert systems designed to compute definite integrals.
SAINT was purely heuristic search based and aimed to be as good as a first year student. SIN combined algebraic manipulation and pattern matching with a search and aimed to be as good as an expert human integrator.

Some History: SAINT and SIN

Some of the first attempts at symbolic integration software were done as Artificial Intelligence Research.
SAINT (1961) by J. Slagle and SIN (1967) by J. Moses were both expert systems designed to compute definite integrals.
SAINT was purely heuristic search based and aimed to be as good as a first year student. SIN combined algebraic manipulation and pattern matching with a search and aimed to be as good as an expert human integrator.
Moses later went on to write the integration engine for Macsyma, the system that would inspire Geddes and Gonnet to start writing Maple in the late 1970s.

Derivative Divides - First Weapon of SIN

Identify the argument of an elementary function that is a factor of the integrand, and check if the derivative of the argument divides the integrand. If so, change variables and integrate.

Derivative Divides - First Weapon of SIN

Identify the argument of an elementary function that is a factor of the integrand, and check if the derivative of the argument divides the integrand. If so, change variables and integrate.

$$\int u'(x)g(u(x))\,dx \rightarrow \int g(y)\,dy$$

Derivative Divides - First Weapon of SIN

Identify the argument of an elementary function that is a factor of the integrand, and check if the derivative of the argument divides the integrand. If so, change variables and integrate.

$$\int u'(x)g(u(x))\,dx \rightarrow \int g(y)\,dy$$

$$\int \frac{e^x}{1+e^x}\,dx = \int 1/y\,dy = \log(y) = \log(1+e^x)$$

Derivative Divides - First Weapon of SIN

Identify the argument of an elementary function that is a factor of the integrand, and check if the derivative of the argument divides the integrand. If so, change variables and integrate.

$$\int u'(x)g(u(x))\,dx \to \int g(y)\,dy$$

$$\int \frac{e^x}{1+e^x}\,dx = \int 1/y\,dy = \log(y) = \log(1+e^x)$$

Most first year calculus integration problems can be solved by this simple method together with expanding and splitting over sums.

The Rest of SIN

1. Matches patterns and uses known formulae to compute integrals.

2. If the integral doesn't match a known pattern, break up using integration by parts or try the Risch algorithm.

The Risch Theorem

In the first half of the 20th century, G.H. Hardy famously claimed "There is no reason to suppose that [an algorithm for indefinite integration of algebraic functions] can be given".

The Risch Theorem

In the first half of the 20th century, G.H. Hardy famously claimed "There is no reason to suppose that [an algorithm for indefinite integration of algebraic functions] can be given".
However, in 1970, R. Risch gave an algorithm to determine if an *elementary function* had an *elementary* indefinite integral, and if so, to compute it*.

The Risch Theorem

In the first half of the 20th century, G.H. Hardy famously claimed "There is no reason to suppose that [an algorithm for indefinite integration of algebraic functions] can be given".
However, in 1970, R. Risch gave an algorithm to determine if an *elementary function* had an *elementary* indefinite integral, and if so, to compute it*.
His method builds on work 19th century work of Louiville and is a generalization of partial fraction algorithm for rational polynomials.

$$\int p(x) + \sum_{i,j} \frac{a_{i,j}}{(x-\alpha_i)^j}\, dx = v(x) + \sum_i c_i \log(u_1(x))$$

The Risch Theorem

In the first half of the 20th century, G.H. Hardy famously claimed "There is no reason to suppose that [an algorithm for indefinite integration of algebraic functions] can be given".
However, in 1970, R. Risch gave an algorithm to determine if an *elementary function* had an *elementary* indefinite integral, and if so, to compute it*.
His method builds on work 19th century work of Louiville and is a generalization of partial fraction algorithm for rational polynomials.

$$\int p(x) + \sum_{i,j} \frac{a_{i,j}}{(x-\alpha_i)^j}\, dx = v(x) + \sum_i c_i \log(u_1(x))$$

Techniques based on Algebra: resultants and differential field theory

Extending Past Elementary Functions

Risch tells us that some integrals cannot be expressed in terms of elementary functions. However, some of these integrals are important enough to have names.

Extending Past Elementary Functions

Risch tells us that some integrals cannot be expressed in terms of elementary functions. However, some of these integrals are important enough to have names. e.g. No closed form for

$$\int e^{-x^2}\,dx$$

But it is an important function, so it has been given a name "erf":

$$\int e^{-x^2}\,dx := \frac{\sqrt{\pi}}{2}\,\mathrm{erf}(x)$$

Extending Past Elementary Functions

Risch tells us that some integrals cannot be expressed in terms of elementary functions. However, some of these integrals are important enough to have names. e.g. No closed form for

$$\int e^{-x^2}\,dx$$

But it is an important function, so it has been given a name "erf":

$$\int e^{-x^2}\,dx := \frac{\sqrt{\pi}}{2}\,\mathrm{erf}(x)$$

Risch can be extended to integrate elementary functions together with erf.

$$\int e^{-(ax^2+2bx+c)}\,dx = 1/2\,\sqrt{\pi/a}\,e^{\frac{b^2-ac}{a}}\,\mathrm{erf}(\sqrt{a}\,x + b/\sqrt{a})$$

Doesn't Risch Solve Indefinite Integration?

In theory Risch solves integration for all functions combining algebraic, exponential, logarithmic, trig and hyperbolic trig functions and their inverses.

Doesn't Risch Solve Indefinite Integration?

In theory Risch solves integration for all functions combining algebraic, exponential, logarithmic, trig and hyperbolic trig functions and their inverses.

But it can produce complicated forms for integrals which may have much simpler forms. We generally want simple forms.

Doesn't Risch Solve Indefinite Integration?

In theory Risch solves integration for all functions combining algebraic, exponential, logarithmic, trig and hyperbolic trig functions and their inverses.

But it can produce complicated forms for integrals which may have much simpler forms. We generally want simple forms.

Many interesting functions are not elementary: e.g. Bessel $J_0(x)$

Doesn't Risch Solve Indefinite Integration?

In theory Risch solves integration for all functions combining algebraic, exponential, logarithmic, trig and hyperbolic trig functions and their inverses.

But it can produce complicated forms for integrals which may have much simpler forms. We generally want simple forms.

Many interesting functions are not elementary: e.g. Bessel $J_0(x)$

In practice, Maple works much like SIN - it analyses the integrand and tries many simple techniques and known formulae before applying the its implementation of the Risch algorithm.

Fundamental Theorem of Calculus (FToC)

As we all learned in first year, the indefinite integral

$$F(x) = \int f(x)\,dx$$

is related to the definite integral

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

Don't forget the fine print: f must be continuous on $[a, b]$
And the missing print: F must also be continuous on $[a, b]$

Algebraic vs. Analytic Indefinite Integrals

Consider the indefinite integral:

$$F = \int 1/(x^2 - 8x + 17)\, dx = \arctan(\frac{x-5}{x-3})$$

Algebraic vs. Analytic Indefinite Integrals

Consider the indefinite integral:

$$F = \int 1/(x^2 - 8x + 17)\,dx = \arctan(\frac{x-5}{x-3})$$

F cannot be used to compute the definite integral:

$$\int_2^4 1/(x^2 - 8x + 17)\,dx = \arctan(2)$$

since F is not continuous on $[2,4]$
(it has a jump discontinuity at $x = 3$)

Algebraic vs. Analytic Indefinite Integrals

Consider the indefinite integral:

$$F = \int 1/(x^2 - 8x + 17)\,dx = \arctan(\frac{x-5}{x-3})$$

F cannot be used to compute the definite integral:

$$\int_2^4 1/(x^2 - 8x + 17)\,dx = \arctan(2)$$

since F is not continuous on $[2,4]$
(it has a jump discontinuity at $x = 3$)
F is an indefinite integral algebraically, but it is not the continuous antiderivative that we know exists from analysis:

$$\hat{F}(x) = \int_0^x 1/(x^2 - 8x + 17)\,dx = \arctan(x-4)$$

Computer Analysis

Since indefinite integrals are computed using algebra, we must be very careful applying the FToC to compute definite integrals.

Computer Analysis

Since indefinite integrals are computed using algebra, we must be very careful applying the FToC to compute definite integrals.

Key Tools:

1. Singularity and discontinuity detection - check if indefinite integrals are continuous
2. Computation of (one-sided) limits - we can use these to "fix" discontinuous integrals and use them for FToC.

Computer Analysis

Since indefinite integrals are computed using algebra, we must be very careful applying the FToC to compute definite integrals.

Key Tools:

1. Singularity and discontinuity detection - check if indefinite integrals are continuous
2. Computation of (one-sided) limits - we can use these to "fix" discontinuous integrals and use them for FToC.

This is can be very hard when the integrand contains symbols other than x.

FToC in Maple

- Compute F, an indefinite integral

FToC in Maple

- Compute F, an indefinite integral
- Call `discont(F)` to determine the discontinuities

FToC in Maple

- Compute F, an indefinite integral
- Call `discont(F)` to determine the discontinuities
- If F has a discontinuity at $x = c$, patch it with limits

$$\hat{F}(x) = \begin{cases} F(x) & x < c \\ F(x) - \lim_{x \to c^+} F(x) + \lim_{x \to c^-} F(x) & c \leq x \end{cases}$$

FToC in Maple

- Compute F, an indefinite integral
- Call `discont(F)` to determine the discontinuities
- If F has a discontinuity at $x = c$, patch it with limits

$$\hat{F}(x) = \begin{cases} F(x) & x < c \\ F(x) - \lim_{x \to c^+} F(x) + \lim_{x \to c^-} F(x) & c \le x \end{cases}$$

- Compute definite integral

$$\int_a^b f(x)\,dx = \lim_{x \to b^-} \hat{F}(x) - \lim_{x \to a^+} \hat{F}(x)$$

FToC Methods in Maple

Maple's `int` command accepts several options:

`int(expr, x=a..b, 'continuous');` avoids checking for discontinuities

`int(expr, x=a..b, 'CauchyPrincipalValue');` computes the limits the the end points, and one-sided limits at singularities simultaneously (cancels infinities)

`int(expr, x=a..b, 'method'='FTOC');` forces only FToC to be tried computing limits with the method by D. Gruntz (i.e. the `limit` command)

`int(expr, x=a..b, 'method'='FTOCMS');` forces only FToC to be tried computing limits with the method by B. Salvy (i.e. the `MultiSeries` package)

Other Methods of Definite Integration

Due to difficulties with FToC and the fact that some integrals do no have closed form indefinite integrals, other techniques are needed.

Most integration engines use pattern matching and heuristic search methods to try to solve many integrals by applying known formulae such as are cataloged in tomes like: Abramowitz & Stegun, Gradshteyn & Ryzhik, and Prudnikov, Brychkov, & Marichev.

In Maple we take care when using these formulae, since these references are known to have many errors.

Elliptic Integrals

All integrals of the form

$$\int_0^a \frac{f_1(x) + f_2(x)\sqrt{T(x)}}{f_3(x) + f_4(x)\sqrt{T(x)}},$$

where the f_i are polynomials, and T is a polynomial of degree at most four, can be written in terms of the three canonical elliptic functions EllipticF, EllipticE, and EllipticPi. e.g.

$$\mathrm{E}(x, k) = \int_0^x \sqrt{\frac{1 - k^2 t^2}{1 - t^2}}\, dt$$

`int(expr, x=a..b, 'method'='Elliptic');`
forces only this technique to be tried.

MeijerG Functions

The Meijer G functions are a broad class of generalized hypergeometric functions

$$G_{p,q}^{m,n}\left(\begin{matrix} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{matrix} \,\middle|\, z \right) = \frac{1}{2\pi i} \int_L \frac{\prod_{j=1}^{m} \Gamma(b_j - s) \prod_{j=1}^{n} \Gamma(1 - a_j + s)}{\prod_{j=m+1}^{q} \Gamma(1 - b_j + s) \prod_{j=n+1}^{p} \Gamma(a_j - s)} z^s \, ds \, .$$

with two useful properties:

1. Elementary and many higher functions can be written as MeijerG functions
2. There are simple formulas for integrals of MeijerG functions and products of MeijerG functions in terms of Γ and other MeijerG functions.

Examples

$$e^x = G^{1,0}_{0,1}\left(\begin{matrix} \\ 0 \end{matrix} \middle| -x \right)$$

$$\arcsin x = \frac{-i}{2\sqrt{\pi}}\; G^{1,2}_{2,2}\left(\begin{matrix} 1, 1 \\ \frac{1}{2}, 0 \end{matrix} \middle| -x^2 \right)$$

$$\ln(1+x) = G^{1,2}_{2,2}\left(\begin{matrix} 1, 1 \\ 1, 0 \end{matrix} \middle| x \right)$$

$$J_\nu(x) = G^{1,0}_{0,2}\left(\begin{matrix} \\ \frac{\nu}{2}, \frac{-\nu}{2} \end{matrix} \middle| \frac{x^2}{4} \right)$$

MeijerG Functions

Integration via MeijerG is very powerful. For example:

$$\int_0^\infty G_{p,q}^{m,n}\left(\begin{array}{c}\mathbf{a}\\ \mathbf{b}\end{array}\middle|\ \eta x\right) G_{\sigma,\tau}^{\mu,\nu}\left(\begin{array}{c}\mathbf{c}\\ \mathbf{d}\end{array}\middle|\ \omega x\right) dx = \frac{1}{\eta}\ G_{q+\sigma,\ p+\tau}^{n+\mu,\ m+\nu}\left(\begin{array}{c}-b_1,\ldots,-b_m,\mathbf{c},-b_{m+1},\ldots,-b_q\\ -a_1,\ldots,-a_n,\mathbf{d},-a_{n+1},\ldots,-a_p\end{array}\middle|\ \frac{\omega}{\eta}\right)$$

MeijerG Functions

Integration via MeijerG is very powerful. For example:

$$\int_0^\infty G_{p,q}^{m,n}\left(\begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\,\middle|\,\eta x\right) G_{\sigma,\tau}^{\mu,\nu}\left(\begin{matrix}\mathbf{c}\\ \mathbf{d}\end{matrix}\,\middle|\,\omega x\right) dx = \frac{1}{\eta}\, G_{q+\sigma,\,p+\tau}^{n+\mu,\,m+\nu}\left(\begin{matrix}-b_1,\ldots,-b_m,\mathbf{c},-b_{m+1},\ldots,-b_q\\ -a_1,\ldots,-a_n,\mathbf{d},-a_{n+1},\ldots,-a_p\end{matrix}\,\middle|\,\frac{\omega}{\eta}\right)$$

However:

1. the conditions under which the formulas hold are complicated
2. conversion to and from MeijerG functions can be difficult

MeijerG Functions

Integration via MeijerG is very powerful. For example:

$$\int_0^\infty G_{p,q}^{\,m,n}\left(\begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\,\middle|\,\eta x\right) G_{\sigma,\tau}^{\,\mu,\nu}\left(\begin{matrix}\mathbf{c}\\ \mathbf{d}\end{matrix}\,\middle|\,\omega x\right) dx = \frac{1}{\eta}\; G_{q+\sigma,\,p+\tau}^{\,n+\mu,\,m+\nu}\left(\begin{matrix}-b_1,\ldots,-b_m,\mathbf{c},-b_{m+1},\ldots,-b_q\\ -a_1,\ldots,-a_n,\mathbf{d},-a_{n+1},\ldots,-a_p\end{matrix}\,\middle|\,\frac{\omega}{\eta}\right)$$

However:

1. the conditions under which the formulas hold are complicated
2. conversion to and from MeijerG functions can be difficult

```
int(expr, x=a..b, 'method'='MeijerG');
```

forces only this technique to be tried.

Definite Integration in Maple

Creating good integration software is about balancing the strengths and weaknesses of the various integration techniques. In Maple 14 it goes something like this:

1. use quick techniques for simple polynomial and rational integrands
2. try to match various formulae, including Elliptic formulae
3. if the integrand has special/higher functions, try MeijerG
4. if not, try FToC using Gruntz limits
5. try MeijerG for any integrand
6. try FToC using MultiSeries limits

Thank you for listening