Symbolic Integration

Victor Adamchik

Carnegie Mellon University

Integrals of a Logarithmic Extension

In this section we outlline the decision procedure algorithm for the log extension defined by

$$D(t) = D(u) / u$$

$$\int \frac{\partial}{\partial x} u(x) \, \frac{dx}{u(x)} = \log u(x)$$

Applying the Euclidean division algorithm and then Hermite's algorithm we obtain

$$\int \frac{a(t)}{b(t)} \, dx = \text{Rational\_Function}\,(t) + \int c(t) + \int \frac{p(t)}{q(t)}$$

where q is squarefree. The first integral is called the polynomial part and the second - the rational part. Unlike to rational function integration, the integration of the polynomial part is the hardest of two.

Polynomial Part

We are to integrate

$$\int c(t)$$

where c is a polynomial in $t(x)$. In other words, c is a polynomial in logs

$$c(t) = a_k \, t^k + \ldots + a_0$$

Integrating this, yields

$$\int c(t) = \int a_k \, t^k + \ldots + a_0 = b_{k+1} \, t^{k+1} + \ldots + b_0 + \sum_{k=1}^{n} d_k * \log\,(u_k)$$

Coefficients $b_p(x)$ can be computed by differentiating both sides wrt x

$$c(\theta) = \frac{d}{d\,x}\left(b_{k+1}(x) \;\; t(x)^{k+1} + \ldots + b_0(x) + \sum_{k=1}^{n} d_k * \log\,(u_k(x))\right)$$

or

$$a_k\, t^k + \ldots + a_0 = b_{k+1}'\, t^{k+1} + \left[(k+1)\, b_{k+1}\, t' + b_k'\right] t^k + \ldots +$$

$$\left[2\, b_2\, t' + b_1'\right]\, t + b_1\, t' + b_0' + \frac{d}{d\,x} \sum_{k=1}^{n} d_k * \log u_k$$

Equating coefficients by t, we get a system of equations

$$\begin{array}{cc} b_{k+1}' & = 0 \\ (k+1)\, b_{k+1}\, t' + b_k' & = a_k \\ \ldots & \ldots \\ 2\, b_2\, t' + b_1' & = a_1 \\ b_1\, t' + b_0' + \frac{d}{d\,x} \sum_{k=1}^{n} d_k * \log u_k & = a_0 \end{array}$$

We solve the system from top to bottom, recursively applying the integration algorithm. Though each step requires integration, the integrand lies in a lower field extension.

The algorithm is terminated when integration cannot be done or the whole system has no solution.

Note, integration itself might require a new LOG extension. In this case we terminate algorithm as well. Though, we could allow a new LOG extension at the last step.

■ Example.

Consider

$$\int \log^2(x) + 2\log(x)\, dx$$

Extending Q[x] by $t = \log(x)$, we obtain

$$\int t^2 + 2\, t\, dx = b_3\, t^3 + b_2\, t^2 + b_1\, t + b_0 + \sum_{k=1}^{n} d_k * \log(u_k(x))$$

where b_k are free of $t = \log(x)$. Differentiating both sides wrt x

$$t^2 + 2\, t = b_3'\, t^3 + 3\, b_3\, t'\, t^2 + b_2'\, t^2 + 2\, b_2\, t'\, t + b_1'\, t + b_1\, t' + b_0' + \frac{d}{d\,x} \sum_{k=1}^{n} d_k * \log(u_k(x))$$

and equating coefficients by t, we get a system of equations

$$\begin{aligned} b_3' &= 0 \\ 3\, b_3\, t' + b_2' &= 1 \\ 2\, b_2\, t' + b_1' &= 2 \\ b_1\, t' + b_0' + \tfrac{d}{d\,x} \textstyle\sum_{k=1}^{n} d_k * \log(u_k) &= 0 \end{aligned}$$

From the first equation

$$b_3(x) = c_3$$

From the second equation

$$b_2' = 1 - 3\, c_3\, t'$$

$$b_2(x) = x - 3\, c_3\, t + c_2$$

Since $t = \log(x)$, we conclude that $c_3 = 0$.

From the third equation

$$2\,(x + c_2)\, t' + b_1' = 2$$

$$2 + 2\, c_2\, t' + b_1' = 2$$

$$2\, c_2\, t' + b_1' = 0$$

$$2\, c_2\, t + b_1 = c_1$$

Since $t = \log(x)$, we conclude that $c_2 = 0$.

From the last equation

$$c_1\, t' + b_0' + \frac{d}{d\,x} \sum_{k=1}^{n} d_k * \log(u_k) = 0$$

$$c_1\, t + b_0 + \sum_{k=1}^{n} d_k * \log(u_k) = c_0$$

Equating coefficients, we obtain $c_1 = 0$ and $b_0 = c_0$ and all $d_k = 0$

Putting all these together

$$\int t^2 + 2\, t\, dx = b_3\, t^3 + b_2\, t^2 + b_1\, t + b_0 + \sum_{k=1}^{n} d_k * \log(u_k)$$

$$\int \log^2(x) + 2\log(x)\, dx = x\, t^2 = x \log^2 x$$

Logarithmic Part

To integrate

$$\int \frac{p(\theta)}{q(\theta)}$$

we use Rothstein-Trager's algorithm:

$$\int \frac{p}{q} = \sum_{k=1}^{n} c_k * \log(d_k)$$

where c_k are the distinct roots of

$$R(z) = \text{res}_\theta(q(\theta),\ \ p(\theta) - z * q'(\theta))$$

and d_k are the polynomials

$$d_k = \text{GCD}(q(\theta),\ \ p(\theta) - c_k * q'(\theta))$$

In a case of transcendental extension, roots of $R(z)$ might depend not only of z but x as well. In the former case, we say that the integral cannot be done in elementary functions.

■ **Example.**

Consider

$$\int \frac{\log(x) - 1}{\log^2(x) - x^2}\, dx$$

We extend $Q(x)$ by $t = \log(x)$. Then the integrand can be rewritten as

$$\frac{\log(x) - 1}{\log^2(x) - x^2} = \frac{t - 1}{t^2 - x^2}$$

Compute a resultant of two polynomials (note that t is a function of x)

$$t^2 - x^2 \ \text{ and } \ t - 1 - z \frac{d}{d\,x}\left(t^2 - x^2\right)$$

Since

$$\frac{d}{d\,x}\left(t(x)^2 - x^2\right) = 2\,t(x)\,t'(x) - 2\,x = \frac{2\,t}{x} - 2\,x$$

we get

```
Resultant[t^2 - x^2, t - 1 - z (2 t/x - 2 x), t] // Simplify
```

$$(-1 + x^2)\ (-1 + 4 z^2)$$

Find roots wrt z

```
Solve[% == 0, z]
```

$$\left\{\left\{z \to -\frac{1}{2}\right\}, \left\{z \to \frac{1}{2}\right\}\right\}$$

Find GCDs:

```
PolynomialGCD[t^2 - x^2, t - 1 - 1/2 (2 t/x - 2 x)]
```

$$1 + \frac{t}{x}$$

```
PolynomialGCD[t^2 - x^2, t - 1 + 1/2 (2 t/x - 2 x)]
```

$$1 - \frac{t}{x}$$

Therefore, the integral is

$$\frac{1}{2}\log\left(\frac{t}{x} + 1\right) - \frac{1}{2}\log\left(1 - \frac{t}{x}\right)$$

or

$$\int \frac{\log(x) - 1}{\log^2(x) - x^2}\, dx = \frac{1}{2}\log\left(\frac{\log(x)}{x} + 1\right) - \frac{1}{2}\log\left(1 - \frac{\log(x)}{x}\right)$$

■ **Example.**

Consider

$$\int \frac{1}{x + \log(x)}\, dx$$

The integrand is $\frac{1}{x+t}$, where $t = \log(x)$ in the field $Q(x, t)$. Compute a resultant of two polynomials

$$t + x \text{ and } 1 - z \frac{d}{dx}(x + t(x))$$

which is

```
Resultant[x + t, 1 - z (1 + 1/x), t]
```

```
-(-x + z + x z)/x
```

Find roots

```
Solve[% == 0, z]
```

```
{{z -> x/(1 + x)}}
```

Roots are NOT free of x, so the integral is not elementary.

Tower of Extensions

Given an integrand

$$f \in K(x, \theta_1, \theta_2, \ldots, \theta_n)$$

where each θ_k is transcendental. The integrand may be manipulated as the rational functions of θ_k. Let us choose the last extension $\theta = \theta_n$

$$f(\theta) = \frac{p(\theta)}{q(\theta)} \in F_{n-1}(\theta) = K(x, \theta_1, \theta_2, \ldots, \theta_{n-1})$$

and integrate f with respect to θ_n. The algorithm is recursive, namely, when we integrate $f(\theta)$ we will be recursively invocate integration in the lower field F_{n-1}.

■ Example

Compute the integral

$$\int\left[\frac{2\log(x+1)}{x}+\log(x)\left(\log(x+1)+\frac{1}{x+1}\right)\right]dx$$

Extensions

$$t_1 = \log(x),\ \ t_2 = \log(x+1)$$

The integral

$$\int \frac{2\,t_2}{x} + t_1\left(t_2 + \frac{1}{x+1}\right)dx$$

It is polynomial in t_1

$$\int \frac{2\,t_2}{x} + t_1\left(t_2 + \frac{1}{x+1}\right) = b_2\ t_1^2 + b_1\,t_1 + b_0 + R \tag{1}$$

where R is free of t_1:

$$R = \sum_{k=1}^{n} d_k * \log\left(u_k\right)$$

We differentiate both sides of (1), to get

$$\frac{2\,t_2}{x} + t_1\left(t_2 + \frac{1}{x+1}\right) = b_2'\,t_1^2 + (b_1' + 2\,b_2\,t_1')\,t_1 + b_0' + b_1\,t_1' + \frac{d}{d\,x}R$$

Equating coefficients by t_1, we get the following system of equations

$$\begin{aligned} b_2' &= 0 \\ 2\,b_2\,t_1' + b_1' &= \tfrac{1}{x+1} + t_2 \\ b_0' + b_1\,t_1' + \tfrac{d}{d\,x}R &= \tfrac{2\,t_2}{x} \end{aligned}$$

From the first equation, we find

$$b_2 = c_2$$

From the second equation, we find

$$2\,c_2\,t_1 + b_1 \quad = \int\left(\tfrac{1}{x+1} + t_2\right)$$

Apply the algorithm recursively (we skip this step) to evaluate the integral on the right side. We obtain

$$2\,c_2\,t_1 + b_1(x) \quad = x\,t_2 - x + 2\,t_2 \ + c_1$$

We conclude

$$c_2 = 0$$

$$b_1(x) = x\, t_2 - x + 2\, t_2 + c_1$$

From the last equation

$$\begin{aligned} b_0' + \ b_1\, t_1' + \tfrac{d}{d\,x} R \quad &= \tfrac{2\,t_2}{x} \\ b_0' + \ \tfrac{x\,t_2 - x + 2\,t_2 + c_1}{x} - \tfrac{2\,t_2}{x} \quad &= -\tfrac{d}{d\,x} R \\ b_0 + \int \frac{t_2\, x - x + c_1}{x} &= -R \end{aligned}$$

Again, we need to apply the algorithm recursively to evaluate the above integral

$$b_0 + t_2\, x - 2\, x + t_2 + t_1\, c_1 = -R$$

Since R is free of t_1, we conclude that $c_1 = 0$. We also set b_0 to c_2.

Summing all the above,

$$\int \frac{2\,t_2}{x} + t_1 \left(t_2 + \frac{1}{x+1} \right) = b_2\ {t_1}^2 + b_1\, t_1 + b_0 + R \tag{2}$$

$$\int \frac{2\,t_2}{x} + t_1 \left(t_2 + \frac{1}{x+1} \right) = (x\, t_2 - x + 2\, t_2)\, t_1 - (t_2\, x - 2\, x + t_2)$$

or

$$\int \frac{2 \log (x+1)}{x} + \log(x) \left(\log(x+1) + \frac{1}{x+1} \right) dx$$

$$=$$

$$(x \log (x+1) - x + 2 \log (x+1)) \log (x) - x \log (x+1) - \log (x+1) + 2\, x$$

References

[1]. M. Bronstein, *Symbolic Integration - Transcendental Functions,* Algorithms and Computations in Mathematics, 2nd edition, Vol 1, Springer-Verlag, 2005