Do Computer Scientists need to "Learn" Calculus of Variations?

In regular differentail calculus courses, one picks a function (which can be thought of as an element of a vector space) and they vary its argument and then see how it changes etc. This is the standard story. But there are other calculus' too. Let us consider a functional which is a map from a vector space V (which in our case will be a function space) to the reals:

$$I : \Phi \in V \to \mathbb{R}, \qquad I[\Phi] = \int F(\Phi(\mathbf{x}), \nabla\Phi, \mathbf{x})\, d\mathbf{x} \tag{1}$$

whwere F is a function of Φ, its derivatives and on $\mathbf{x}$. But note I is not a function of $\mathbf{x}$.

We will hereafter think that the change of Φ is caused by the change of its functional form but not in its argument $\mathbf{x}$. Let me give an example by considering a simple example in 1D. Say $\Phi(\mathbf{x})$, as its basis, has $1, x, x^2$, so that

$$\Phi(x) = \alpha 1 + \beta x + \gamma x^2 \tag{2}$$

So the change in the function (in this view) is caused by the change in the components $\{\alpha, \beta, \gamma\}$.

Therefore, when the basis (2) is used in the example (1) one should get

$$I = I(\alpha, \beta, \gamma)$$

Therefore after a choice of finite number of "basis" for the function, the problem will reduce to a problem of multivariate calculus.

Let us look at a more standard example, the problem of "harmonic maps". The"energy" functional for these type of maps is given by

$$I = \frac{1}{2} \int d^D\mathbf{x}\, [\ \nabla\Phi(\mathbf{x}) \cdot \nabla\Phi(\mathbf{x})\ ]\,. \tag{3}$$

If the function I is an extremum, its change under an infinitesimal change in Φ should be zero (just like the extrema of a function):

$$I[\Phi + \delta\Phi] - I[\Phi] \equiv E[\Phi] \cdot \delta\Phi = 0 \Rightarrow E[\Phi] = 0 \tag{4}$$

for infinitesimal $\delta\Phi$, where we have tossed out terms quadratic and higher in $\delta\Phi$. For example (2), this change will be caused by the changes $(\delta\alpha, \delta\beta, \delta\gamma)$ i.e. for each component of the vector $\delta\Phi$ which are all infinitesimals of first order.

Let us compute the variation when one has (3) for our system (which we take to be one dimensional) :

$$I[\Phi + \delta\Phi] - I[\Phi] \equiv \delta I = \frac{1}{2} \int dx\ \left[\left(\frac{d\Phi}{dx} + \frac{d(\delta\Phi)}{dx} \right)^2 - \left(\frac{d\Phi}{dx} \right)^2 \right] = \int dx\ \frac{d\Phi}{dx} \frac{d(\delta\Phi)}{dx}$$

where we have dropped the term quadratic in $\delta\Phi$ (rather its derivative).

Integrating by parts and dropping the endpoint terms (which are zero as $\delta\Phi$ is zero at those endpoints), one gets

$$\delta I = -\int dx\ \frac{d^2\Phi}{dx^2}\delta\Phi$$

As $\delta\Phi$ is arbitrary everywhere along the curve and this is true for any infinitesimal $\delta\Phi$ around the "stationary" point, we are then led to the equation for the extrema:

$$\frac{d^2\Phi}{dx^2} = 0 \tag{5}$$

which is nothing but the Euler-Lagrange equation for this system.

Now I am going to claim that if the discretized theory has the right symmetries as the original theory - the ensuing relation will be the correct finite difference equation which will give the same equation in the continuum limit.

Noet that the functional (3) is invariant under $x \leftrightarrow -x$ and just simply means the energy is independent of the orientation of the line. This feature must be respected in the discrete version, so that the derivative must be replaced by left and right finite differences at each point labeled by x_i:

$$I[\{\Phi_i\}] = \sum_i \left(\frac{\Phi_{i+1}-\Phi_i}{\Delta x}\right)\left(\frac{\Phi_i-\Phi_{i-1}}{\Delta x}\right) = \frac{1}{(\Delta x)^2}\sum_i \Phi_i\Big[\,(\Phi_{i+1}+\Phi_{i-1})-\Phi_i\Big]$$

As before we vary Φ_i independently (i.e. pointwise)

$$\delta I = \sum_i \delta\Phi_i\Big[\frac{(\Phi_{i+1}+\Phi_{i-1})-2\Phi_i}{(\Delta x)^2}\Big]$$

As $\delta I = 0$ under $\delta\Phi_i$ which are (linearly) independent one must have the coefficients of $\delta\Phi_i$ zero for each i , leading to

$$\Big[\frac{(\Phi_{i+1}+\Phi_{i-1})-2\Phi_i}{(\Delta x)^2}\Big]$$

which is nothing but the discrete analog of (5).

So the apparent lesson is that if we understand the correct finitary approximation to the continuum theory, one just can use the standard multivariate calculus techniques without recoursing to Euler-Lagrange eqs. or some such things.