Kelly's optimal position sizing results are practical and elegant. It answers the question, "What fraction of my money should I bet on each flip of a sequence of biased coin flips to maximize my long-term wealth?"

The setup and proof are simple,

Let $W_t$ be your total wealth after $t$ bets, let $\not{f}$ be the fraction, between $[0,1]$ of your wealth/capital you wager on the $t$-th bet, and let $\boldsymbol{\mathcal{R}}$ be the result of each coin flip. Working backwards we can easily get an equation for $W_t$

$$W_t = W_{t-1} + (\not{f} * W_{t-1})\boldsymbol{\mathcal{R}}_t$$

i.e. Your current money is equal to the money you had last bet, $W_{t-1}$, plus the wager, $(\not{f} * W_{t-1})$, times the outcome, $\boldsymbol{\mathcal{R}}_t$. $\boldsymbol{\mathcal{R}}_t = 1$ if you win and $\boldsymbol{\mathcal{R}}_t = -1$ if you lose. Rewriting this via algebra,

$$W_t = W_{t-1} + (\not{f} * W_{t-1})\boldsymbol{\mathcal{R}}_t = W_{t-1}(1 + \not{f} * \boldsymbol{\mathcal{R}}_t)$$

and finally working backwards all the way we extend the equation again,

$$W_t = W_{t-1} + (\not{f}_t * W_{t-1})\boldsymbol{\mathcal{R}}_t = W_{t-1}(1 + \not{f}_t * \boldsymbol{\mathcal{R}}_t) = W_0 \prod_{\tau=1}^{t}(1 + \not{f}_\tau * \boldsymbol{\mathcal{R}}_\tau)$$

For the less mathematically inclined, $\prod_{\tau=1}^{t}(1 + \not{f}_\tau * \boldsymbol{\mathcal{R}}_\tau)$ is a concise way of writing the total compounded return. So your current wealth, $W_t$, is equal to your starting wealth, $W_0$, times your compounded return.

If you don't like math, you probably don't like logarithms, but we need them.

Our goal is obviously to maximize $W_t$, and to do this we need to define a new quantity that's easier to work with, namely the geometric growth rate of our wealth when following the betting strategy $\not{f}$, which we will call $G_t(\not{f})$. Maximizing $G_t(\not{f})$ is equivalent to maximizing $W_t$.

$$G_t(\not{f}) \stackrel{\text{def}}{=} \frac{1}{t}\ln\left(\frac{W_t}{W_0}\right) = \frac{1}{t}\sum_{\tau=1}^{t}\ln\,(1 + \not{f} * \boldsymbol{\mathcal{R}}_\tau)$$

Now we have a nice sum to work with instead of a nasty product. By the law of large numbers/Central Limit theorem,

$$G(\not{f}) = \lim_{t\to\infty}\frac{1}{t}\sum_{\tau=1}^{t}\ln\,(1 + \not{f} * \boldsymbol{\mathcal{R}}_\tau) \;=\; \mathrm{E}[\ln(1 + \not{f} * \boldsymbol{\mathcal{R}})]$$

Where the expectation only applies to the random variable $\boldsymbol{\mathcal{R}}$, which we assume is i.i.d. We'll introduce one last variable here, $\rho$, a little greek p standing for the probability of winning the bet, i.e. the coin landing on heads.

$$\rho \stackrel{\text{def}}{=} P(\boldsymbol{\mathcal{R}}_t = 1)$$

$$= 1 - P(\boldsymbol{\mathcal{R}}_t = -1)$$

So

$G(\not{f}) = \mathrm{E}[\ln(1 + \not{f} * \boldsymbol{\mathcal{R}})] = \rho \ln(1 + \not{f} * 1) + (1 - \rho) \ln(1 + \not{f} * -1)$

As you learned in calculus 1a, to optimize $G_t(\not{f})$, we take a derivative with respect to $\not{f}$ and set the result equal to zero. The derivative of $\ln(x) = \frac{1}{x}$, so we solve the following for $\not{f}$,

$$0 = \rho \frac{1}{1+\not{f}}(1) + (1 - \rho) \frac{1}{1-\not{f}}(-1)$$

$$= \frac{\rho(1-\not{f}) - (1-p)(1+\not{f})}{(1+\not{f})(1-\not{f})}$$

$$= \frac{-1+2\rho-\not{f}}{(1+\not{f})(1-\not{f})}$$

And since we don't care about the denominator unless $\not{f} = \{1\ or - 1\}$, we can finally solve

$$0 = -1 + 2\rho - \not{f}$$

$$\Rightarrow \not{f} = 2\rho - 1$$

Which is the fabled **Kelly Fraction**, the optimal fixed fraction of your wealth to wager on each bet [i.e. trade]. Basically, it is best to bet two times the odds of winning, minus one. So for ex. if you have 65% accuracy and you win the same amount on each winning bet as you would lose on a losing bet, then you should wager $2 * .65 - 1 = 1.3 - 1 = .3 = 30\%$ of your capital on each bet to maximize long-term wealth.

Just for kicks, if we plug this in to $G(\not{f})$ to see how rapidly we'll get rich, we find,

$$G(\not{f}) = \ln(2) + \rho \ln(\rho) + (1 - \rho) \ln(1 - \rho)$$

And here's a very deep fact,

$$\ln(2) + \rho \ln(\rho) + (1 - \rho) \ln(1 - \rho) = \ln(2) - H(\boldsymbol{\mathcal{R}})$$

Where $H(\boldsymbol{\mathcal{R}})$ is the measure of "entropy" of the distribution of $\boldsymbol{\mathcal{R}}$ from information theory literature. If you want to take this analysis beyond the simple case of a two outcome betting game with equal payouts, you will need the tools of information theory (or advanced stochastic calculus). Our present result showing that $\not{f}_{optimal} = 2\rho - 1$ is actually very useful if you think of moving to shorter and shorter timeframes. Then the binary outcomes become a continuous distribution and the conclusions seem a lot more applicable.

The problem we've solved here is the same problem of telling a computer how much to trade. Not only must it predict well, it should also size positions intelligently.

Finally, I need to mention that if your estimate of $\rho$ is inaccurate, i.e. your trading system's accuracy varies, then it will result in incorrectly sized bets, and lost profits. And it turns out that the penalty for betting too large is worse than the penalty for betting too small. So, if in doubt, be conservative.