A Sigmoid Dialogue

By Anders Sandberg

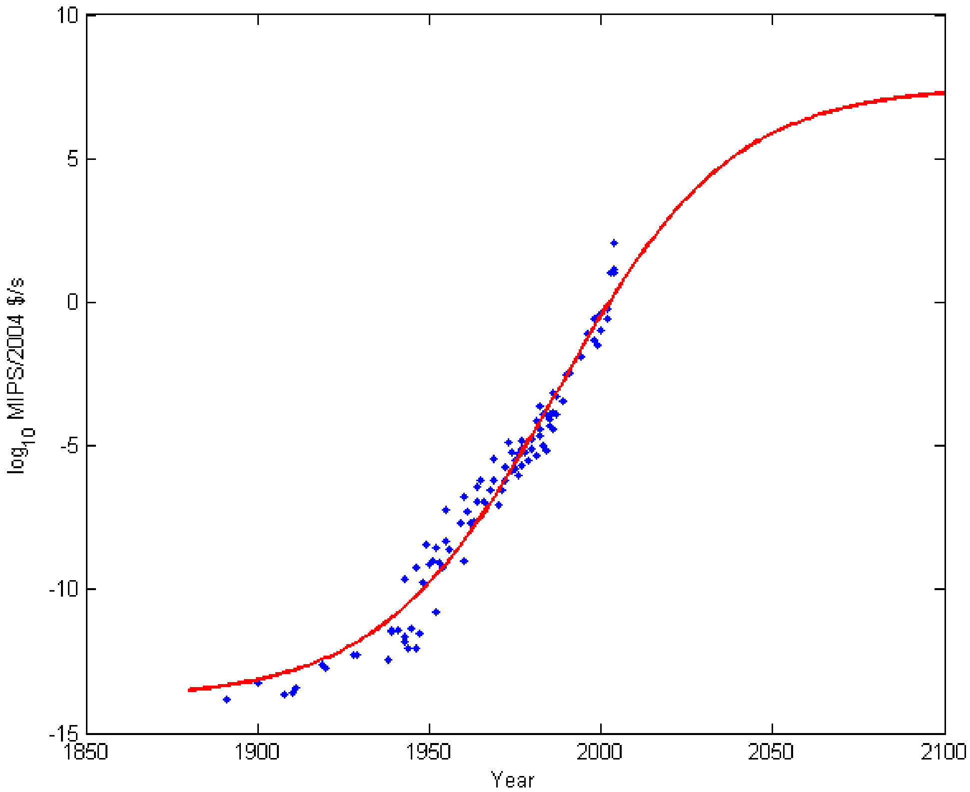
Achilles: Now I know how powerful computers are going to become!

Tortoise: How?

Achilles: I did curve fitting to Moore's law. I know you are going to object that technological progress cannot be exponential forever, so be assured I did a conservative analysis by fitting a logistic curve.

Tortoise: So you assume that over time, like many other technologies, computer performance will first accelerate until some limitations make it slow down and eventually approach some ultimate limit?

Achilles: Exactly. And that limit is my estimate of how powerful computers will ever become. I used a least squares fit using Matlab's fmincon function to get the most likely sigmoid. The data is from the performance curve database http://pcdb.santafe.edu/. Here is my graph:



And as you can see, we will get $10^{7.5}$ flops per dollar in the long run. That is more than 300,000 times our current best performance!

Tortoise: Hmm.

Achilles: You are not convinced, my friend?

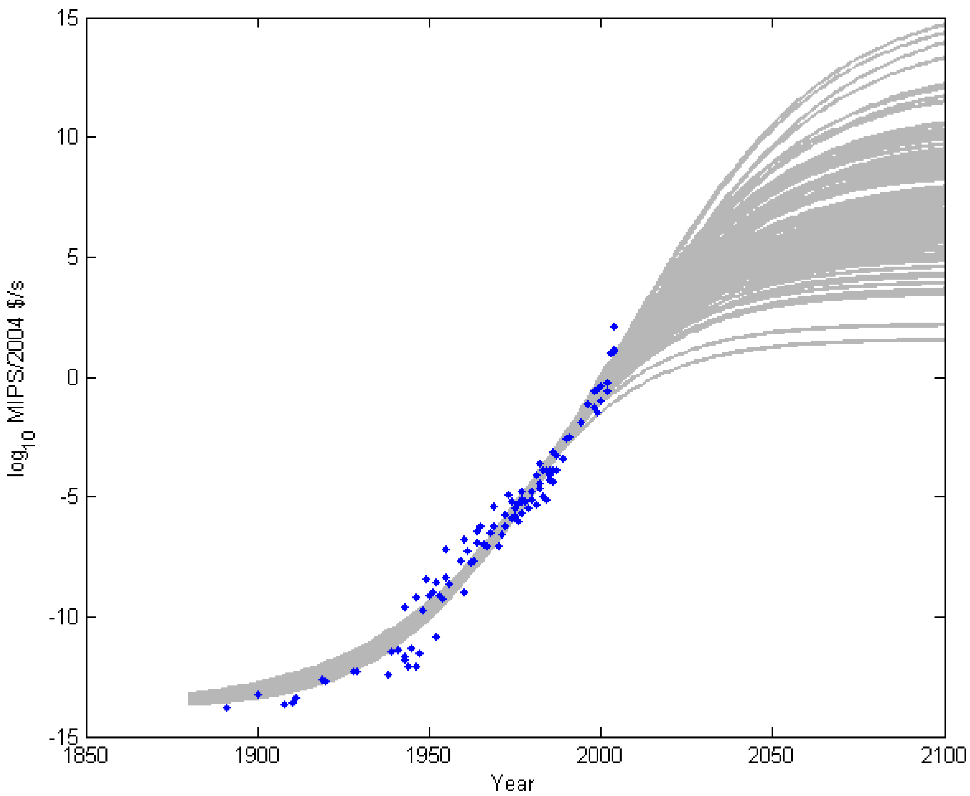
Tortoise: No. How can you be certain of this result, since the data points appear to be a bit random?

Achilles: There is always noise in data; we have to do the best we can with the numbers we got. In the future we can always update this forecast.

Tortoise: As more data arrives your estimate will change, but will it change a little or a lot?

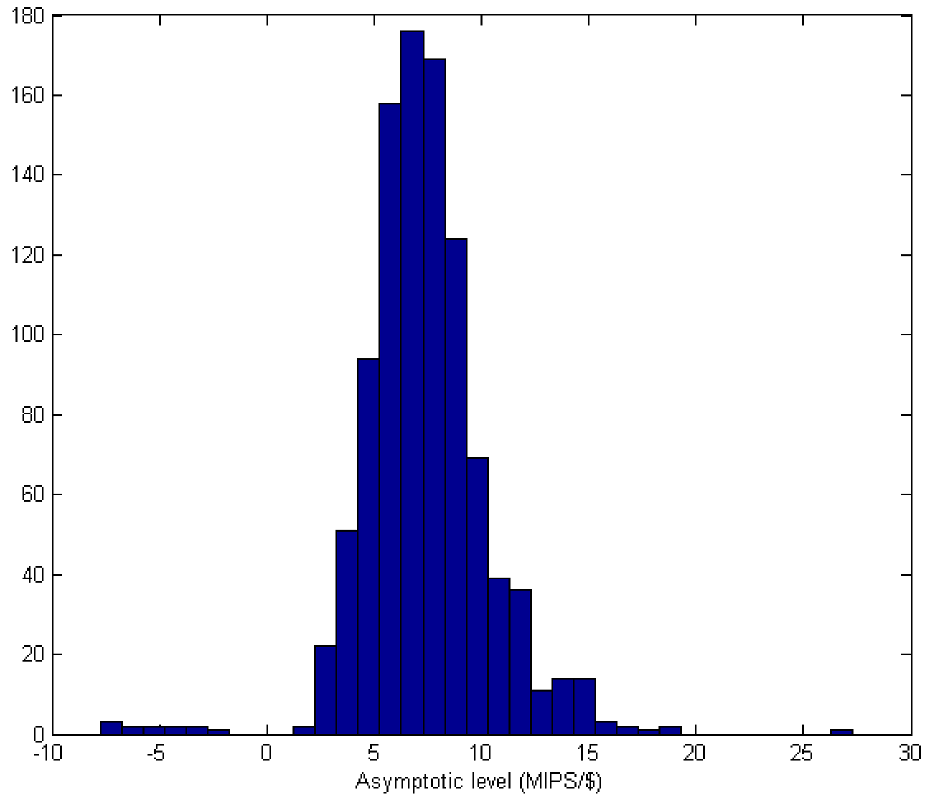
Achilles: OK, that is a problem. But we can do a bootstrapping confidence interval. I resample (with replacement) data points, fit the curve to this proxy dataset, repeat this a lot of times and get a probability distribution of the parameters. Including the ultimate limit. Then I just find the interval containing 90% of the results, and I will have a good estimate of the confidence interval.

Here it is:



And here is a histogram of the parameter:

:



The 95% confidence interval for the ultimate level is from $10^{2.96}$ to $10^{13.99}$. Happy?

Tortoise: Hmm. 11 orders of magnitude of uncertainty in eventual computer power.

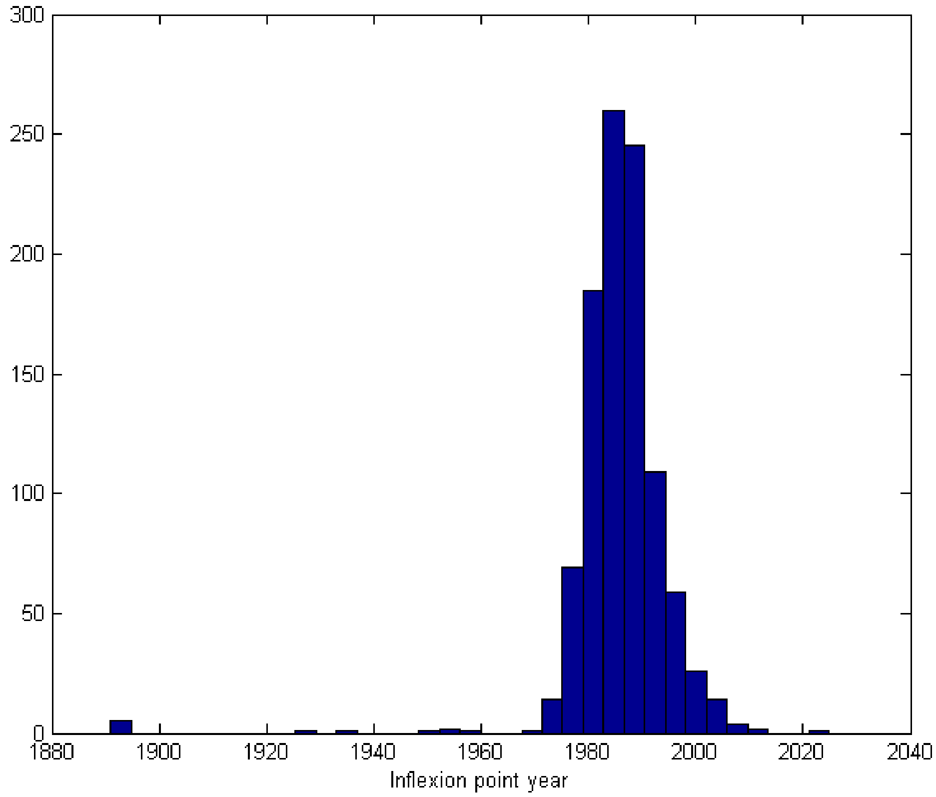
Achilles: Well, no forecast is perfect. Still, this gives us **some** useful information about the ultimate limits.

Tortoise: Maybe. If I were to make a forecast, I would be most concerned with what we could see in the near future. That would allow me to check whether I was likely right or wrong.

Achilles: Growth sigmoids have three parameters: how far up they will go, how sharply they grow when they grow, and when the growth is fastest (i.e. its inflection point). You are suspicious about the first, the second might be more to your liking and I think the third is exactly what you look for.

Tortoise: Yes, knowing how far we have to go has a great deal of practical importance. And when previously accelerating growth starts to slow we can actually notice it a short while afterwards. Could you tell me when Moore's law will have its inflection point?

Achilles: Sure. It will occur in… 1986. Well, with 95% confidence between 1975 and 2002.



Tortoise: Interesting “prediction”.

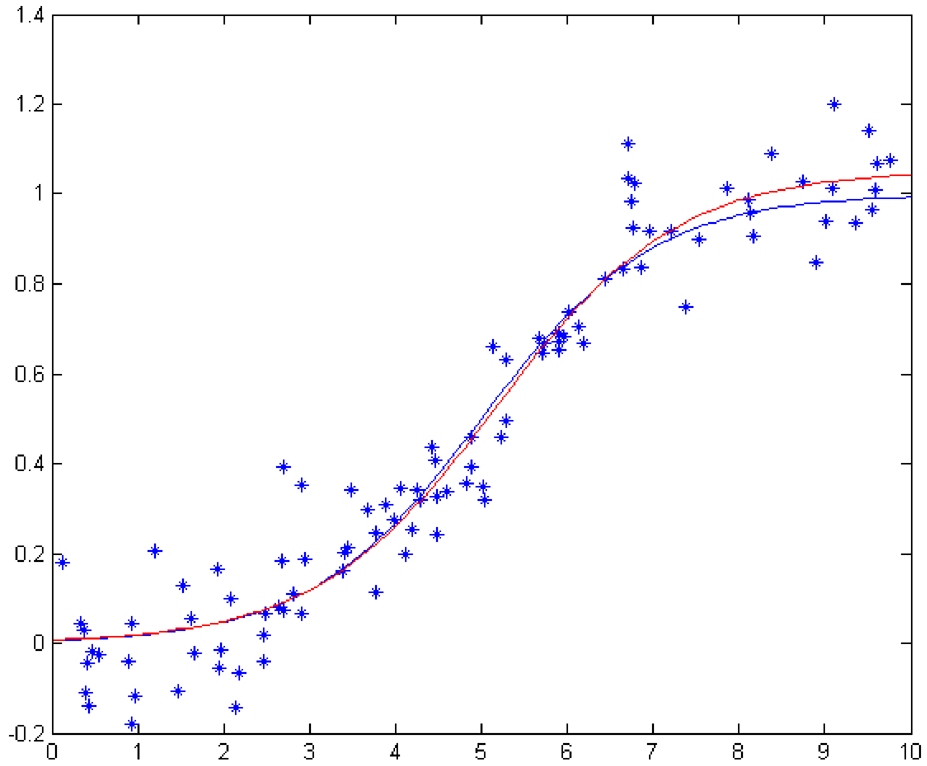
Achilles: Well, I don’t mind being called conservative.

Tortoise: So you trust these estimates?

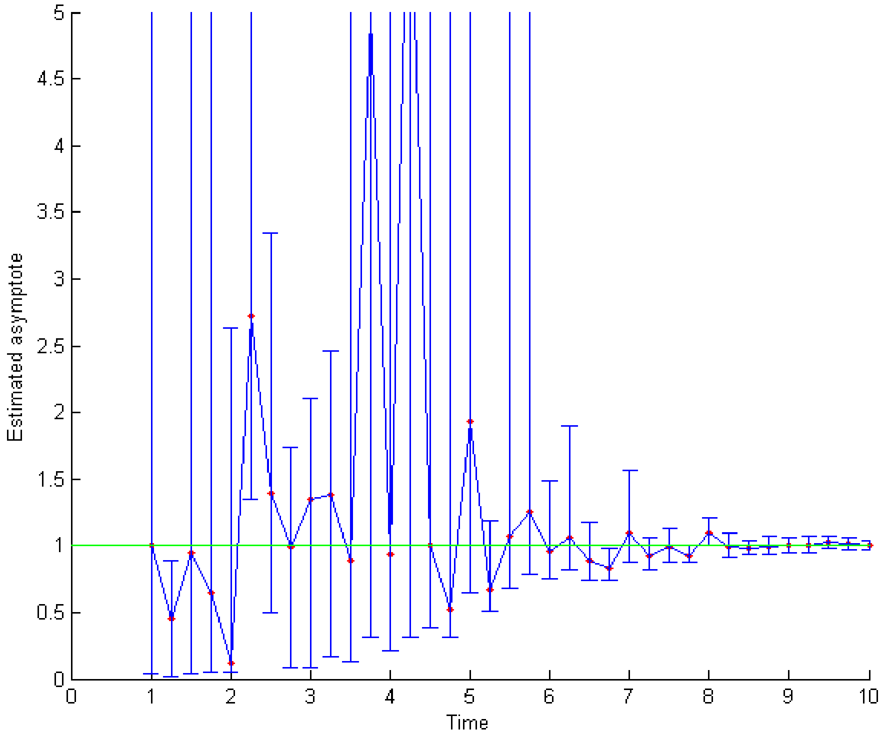
Achilles: Sure. Although it was a bit worrying to see how much uncertainty is hidden in the estimates. But I guess this is natural, it is still early days even if we have just passed the inflection point.

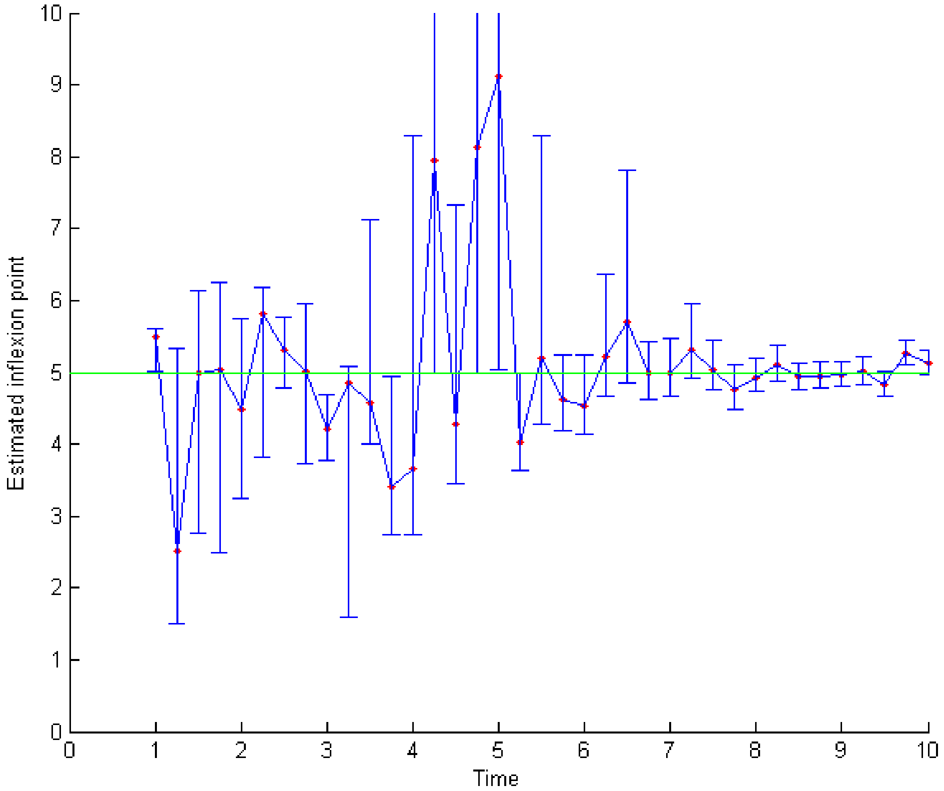
Tortoise: I wonder how reliable predictions like this are before the inflection point. In fact, can you reliably predict when inflection points occur before they happen?

Achilles: I can create some artificial data and test it. Here is a sigmoid with inflection point at time 5, asymptote 1 and sharpness 1, where there are 100 data points with N(0,0.1) noise added. The red curve is the fit, the blue curve is the “real” curve.



OK, let's try this kind of fit when the 100 data points lie in the interval 0 to T for increasing T. Also, I will use 1000 bootstrap samples to estimate the asymptote and when the inflection point will happen. Here are the estimates with confidence intervals:





Oh dear…

Tortoise: So we should not trust the estimates at all until well *after* the inflection point.

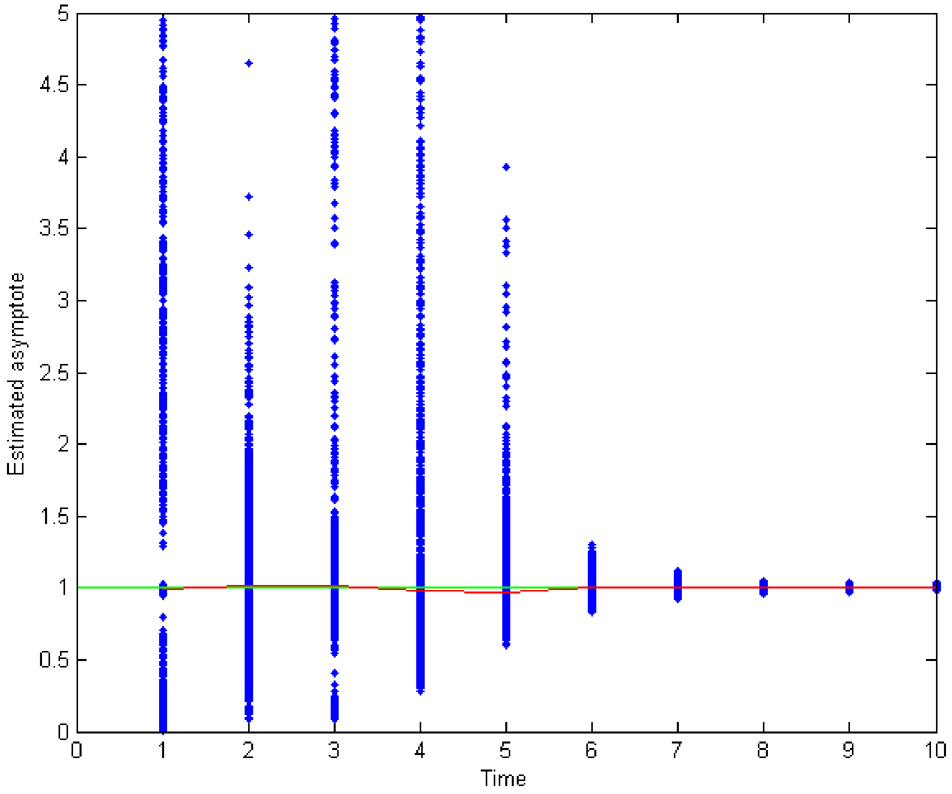
Achilles: Including our estimates of when it will happen or has happened.

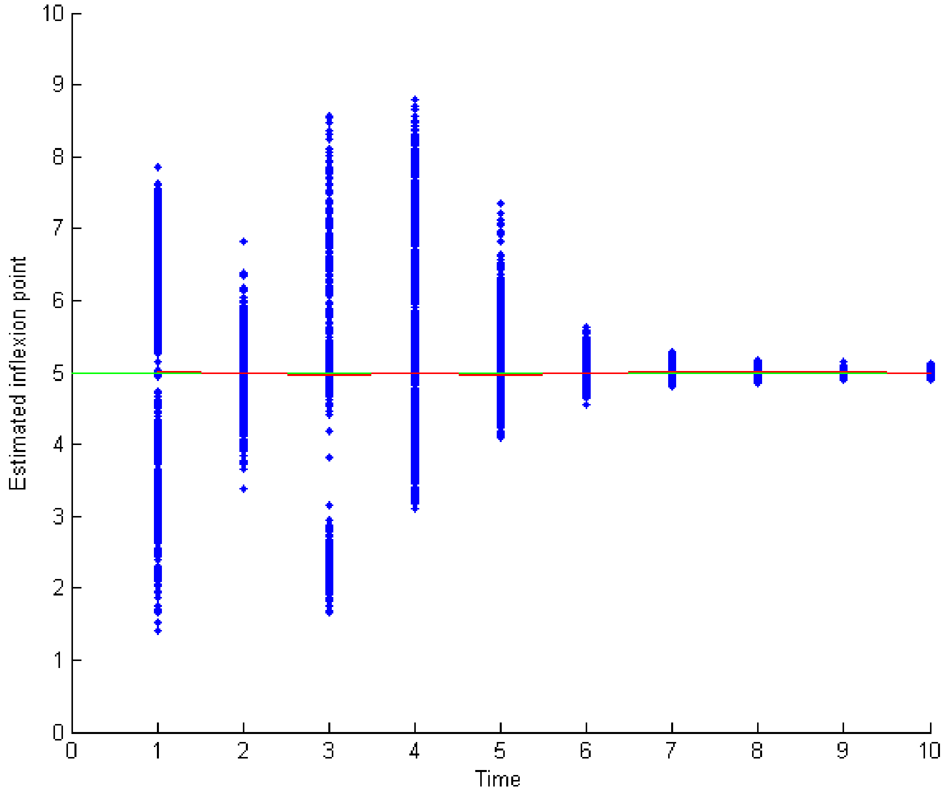
Tortoise: This makes me doubt the claim we have passed the inflection point. Either we have, and we have an accurate estimate, or we have not, and we just got a random estimate.

Achilles: Yes. But these curves are a bit unnecessarily nasty, since each plotted point corresponds to a different trial with different random noise in the given data. Sometimes the data forms a really misleading pattern, but not always.

Tortoise: But we cannot tell the difference when we look at our data. Also, I suspect that for sharper transitions the predictions are even worse before they happen. By the way, I wonder whether there is a bias towards overestimating the eventual level and when the inflection point happens in the middle of the curve?

Achilles: Let me run a lot of independent trials and see. Each dot is a trial, the red line is the median estimate they make.





OK, the median is not bad at estimating how the curve looks. At least it is not biased.

Tortoise: But in real life we cannot get it, since we will only have **one** of those blue points. And at least the eventual level has a very skew distribution.

Achilles: It is interesting to see that the uncertainty at time 1, 3, 4, and 5 is a bit larger than the one at time 2.

Tortoise: I guess that is because it is really hard to predict any shape when the stretch of data is shorter than the spread due to noise, and tricky to do it while the change is maximal. But at time 2 there is at least a detectable straight line. All in all, it shows just how unpredictable this kind of forecasting is.

Achilles: So you think computers are not going to develop as far as I do?

Tortoise: Not at all. I am very optimistic about how far things can go. I just think we shouldn't trust this kind of extrapolation too much. Even if it is true technology progresses like a sigmoid, we cannot reliably use it to predict how far we are going until it is obvious anyway. And who knows, maybe the true curve is really an infinite exponential or an asymmetric sigmoid-like curve (like the solution to $Y'(t)= Y^2(t) - Y^3(t)$). But these models have the same problem. The more complex the curve, the more uncertain the predictions will be.

Achilles: So we should use straight lines instead?

Tortoise: They don't make much sense for the problem given what we actually know about technology growth. Prior information is always important. You used constraints on the sigmoid curve parameters to keep them positive, and this no doubt improved reliability. We should do what we can to make plausible and robust predictions – but we shouldn't think they are going to be correct even if we have the right model.

Achilles: Want to bet that that people will prefer objective-sounding numbers anyway?

Tortoise: That sounds like a robust and plausible prediction.