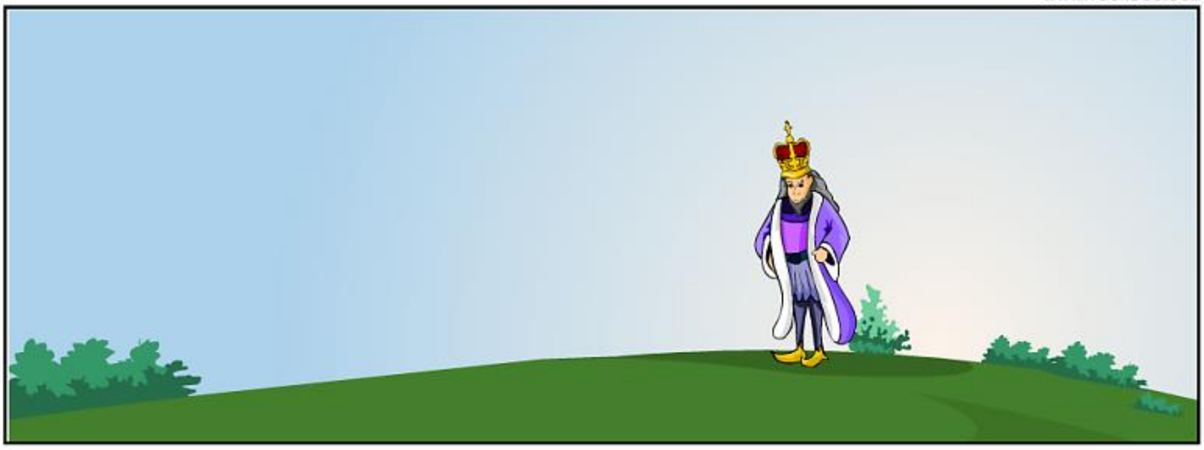
Prologue: A problem for a Particle Filter

Let us take as our example the task of trying to forecast temperatures for a mountain hilltop.

In the distant land of Assimilaris the King has chosen to live on top of a mountain. Because of the position of this mountain and the curious winds that blow over the kingdom of Assimilaris, temperatures on the top of the mountain swing randomly from warm to cold with chaotic irregularity.



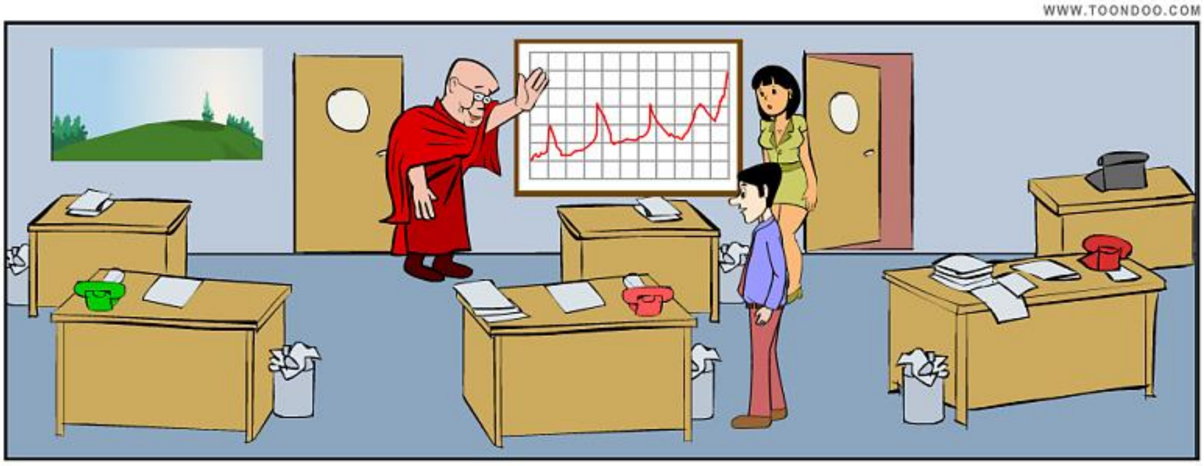
The king, having lived up on this mountain for most of his life, has measured these temperatures for years and found that they can be described by the following equation:

$$\boldsymbol{x^{n+1} = 0.5x^n + 25\left(\frac{x^n}{1+(x^n)^2}\right) + 8cos(0.8n) + \beta^n}$$

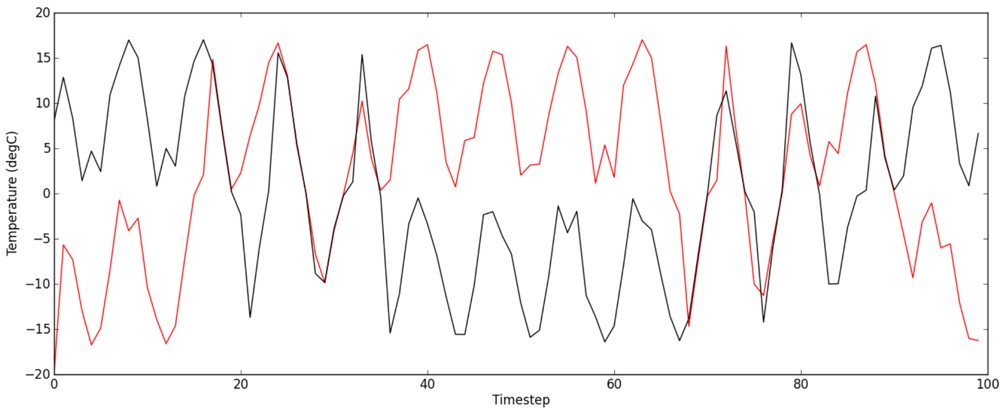
where $\boldsymbol{\beta}^n$ is a small amount of random error that he hasn't yet got to the bottom of.

Now though he has chosen to live away from them, the king's subjects have great admiration for their sovereign and showering him with gifts is the national pastime. The only gifts that the peculiar king is interested in, however, are ice cream or hot chocolate. He also cannot bear waste and so when his subjects bring him ice cream on a cold day - which obviously makes it inedible - then he can get very angry indeed.

So important is this national pastime to the people that the Royal Department of Gifts was set up in order to assist the people in pleasing their King. In the department, the Gift Advisers work constantly on trying to perfect their predictions of the weather on the mountain top on any given day so that they can tell the people what to carry up the mountain to the king.



Once the King had created his equation, the relieved Gift Advisers set about the process of solving the equation. Believing this would lead them to be able to predict the temperatures perfectly. But in fact, it led to disaster:



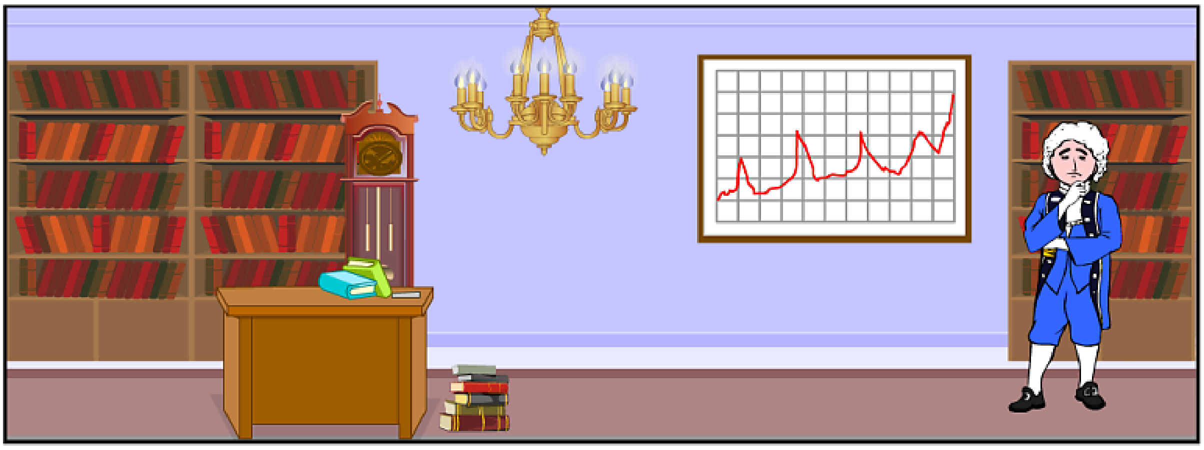
In the image above, the black line is the temperature which the King felt and the red line is the prediction of the Gift Advisers.

The advisers quickly realised that the random error $\boldsymbol{\beta}^n$ in the King's Equation, meant that the system was truly chaotic and the smallest difference between their estimate of the error and the true value would have a huge effect on the calculated temperature change. It was therefore impossible for the Gift Advisers to pin down the temperature this way.

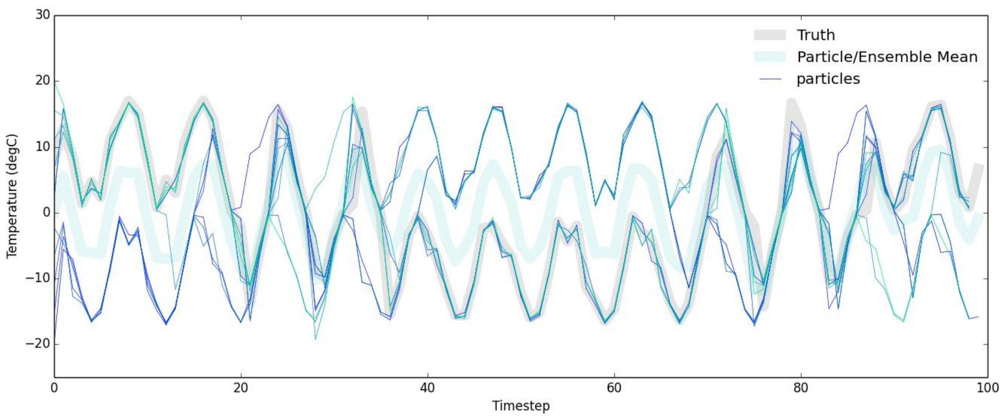
The king was absolutely furious after being inundated with ice-creams during a spate of cold weather. When they came up to the mountain to beg his forgiveness, the King demanded that the Gift Advisers employ a specialist to solve the problem.

Model Ensembles

After being given a thorough dressing down by the King, the advisers created the post of Head Gift Forecaster. After hearing about his excellent work, they gave the post to the distinguished specialist, Lord Ensemble.



Lord Ensemble was convinced that using a number of different versions of the model in an ensemble would be their solution and - to make his mark - he decided to call each of his models *a particle*. He began his work by running an ensemble of 10 particles and then looked at his results:



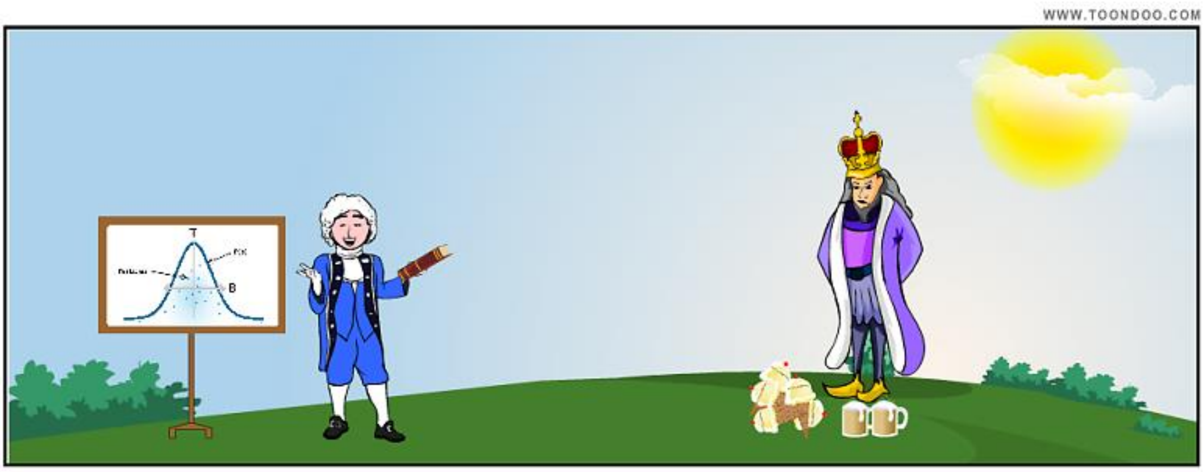
This image shows the truth in a thick grey line and all of the particles in the thin dark lines which are all different shades of blue. The thick faint blue line shows the average of all the particles - this was Lord Ensemble's best guess at the truth.

Run the ensembles python code (see downloadable link) to find out the spread of particles from a different angle. What Lord Ensemble saw is that the black star is the same as the thick grey line and shows the location of the truth. Each of the coloured blocks represents one particle and individual particles can be followed by tracking the colour of the blocks.

Lord Ensemble's method of forecasting the King's weather for the people of Assimilaris was certainly a little better than the single prediction previously used by the Gift Advisers. Ultimately, however, the King came to realise that Lord Ensemble's best guess always floated somewhere in between the two extremes of temperature. This meant that the King would only get a trickle of ice cream or hot chocolate on days when the conditions were extreme enough to warrant extra helpings.

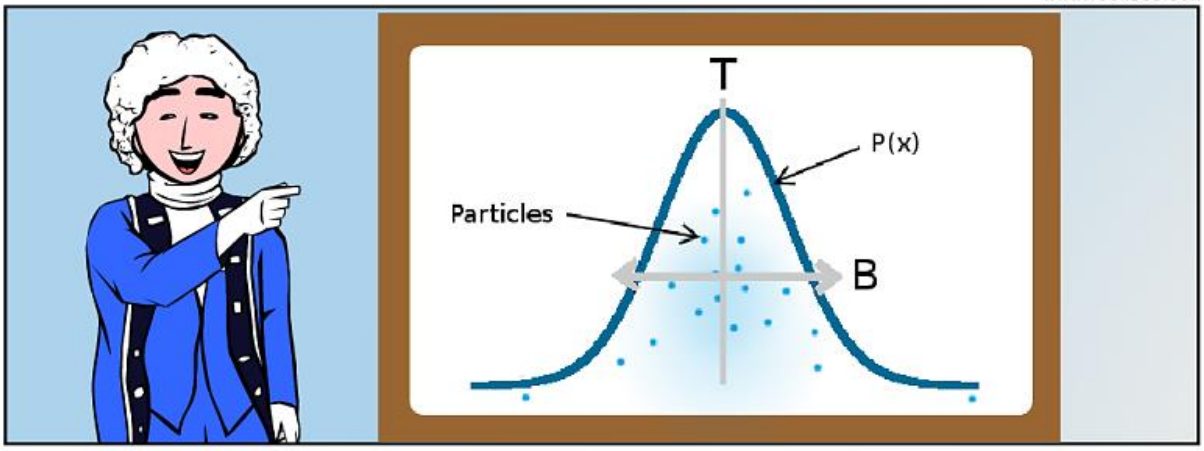
Though the King had a fondness for Lord Ensemble - he could not simply accept that this was the best that could be done. He commanded Lord Ensemble to the mountaintop to explain his mathematics and see if anything could be done to improve the situation.

Ensemble Forecast Theory



Up on the mountain, Lord Ensemble explained his particle ensembles to the King. Under questioning he revealed that his method implicitly assumed that most of the particles he created using the king's equation would lie close to the truth.

Lord Ensemble showed that the position of any given particle (**x**) could be incorporated into a distribution of all particles (**p(x)**) and he believed that this would have a regular and predictable shape around the truth (**T**). Lord Ensemble showed that he believed the particles would have a Normal Distribution, where the mean of particles was around the truth.



ord Ensemble explained that the rate at which particles numbers decreased with distance from the peak was the variance of the distribution and that this could be quantified. The smaller this was, the closer the particles would generally be to the truth. He called the variance **B**.

Using these definitions, Lord Ensemble showed the king that the distribution of the particles would therefore give him information about the value of **T**:

$$p(x) \approx N(T, B)$$

Lord Ensemble was clearly proud of his work and so it was with some regret that the King pointed out to him that the particles *didn't* have a normal distribution around the truth. The chaotic mountaintop temperatures tended to cluster in warm spells and icy spells and spent very little time in-between at freezing point. Therefore, the King explained, his approach was far too simplistic.

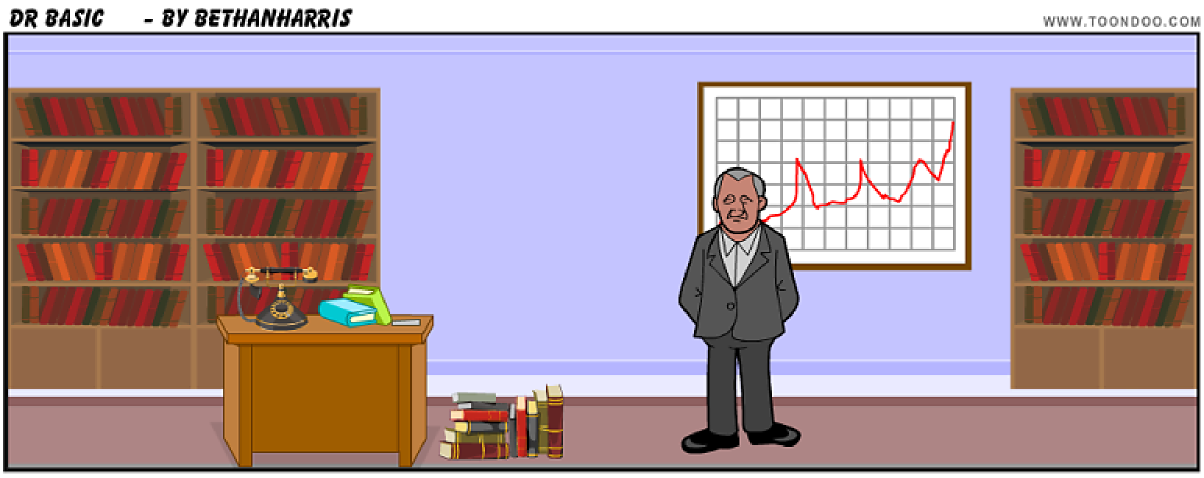
The humbled Lord Ensemble saw his error and promised the King that he would try to find the right probability distribution to give the best estimate of the truth.

Lord Ensemble rushed back to the palaces and spent months desperately trying to find a distribution that would work perfectly. Despite all his efforts, however, he could not find a satisfactory solution. He became irritable and unpleasant to works with. The Gift Advisers, tried to reason with him, to suggest that he try a different course of attack, but Lord Ensemble refused to be shaken from his quest.

Thus is was that - after a quiet word from the Gift Advisers - Lord Ensemble was once again summoned to the mountaintop by the King. This time to be reluctantly relieved of his position on the grounds of standing in the way of progress.

Basic Particle Filtering - Observations

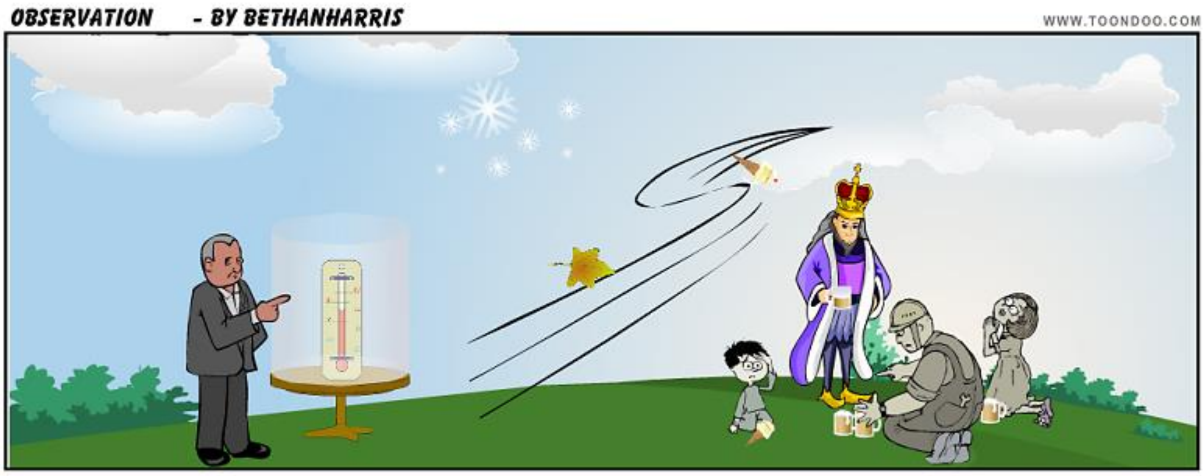
After being given yet another thorough dressing down by the King for their poor recruitment techniques, the Gift Advisers embarked on a very careful process of checks and interviews in order to fill the now vacant position of Head Gift Forecaster. From the small handful of candidates that applied for the position, then eventually gave the job to a promising academic called Dr Basic.



Dr Basic would eventually go down in the history books of Assimilaris as the inventor of the first ever particle filter. He took the foundation laid down by Lord Ensemble and advanced it with a significant development - incorporating observations.

Dr Basic realised very early on that the chaotic nature of the mountain winds was such that it was always impossible to accurately guess the mountain top temperature from the King's Equation alone. Not wishing to repeat Lord Ensemble's mistake, Dr Basic decided to try a completely different approach.

Dr Basic thought that taking observations of the actual temperature would allow them to give the best guidance to the people of Assimilaris. He spoke to the King who agreed to allow him to come up to the top of the mountain in order to take a reading of the temperature. He was, however, required to use the Royal Thermometer which had been kept safely in a kitchen cupboard for nearly 300 years.



Once he had his observations, rather than merely trusting the thermometer to tell him the absolute truth, Dr Basic did some research and found that the readings taken from the rusty old thermometer had some error on them. Referring to the dusty old user's manual which came with the thermometer - he knew that that error was equal to **R**.

The old manual also told Dr Basic that these errors were normally distributed around the Truth at the time that the observation was made. So each observation (**y**) was a sample from a normal distribution given the Truth:

$$p(y|\mathrm{T})=\mathrm{N}(\mathrm{T},\mathrm{R})$$

But unfortunately it quickly became apparent that observations on their own weren't enough. Dr Basic could only face the walk up the mountain once a week or so and the observations on their own told the Gift Advisers nothing about what was happening in between observations.

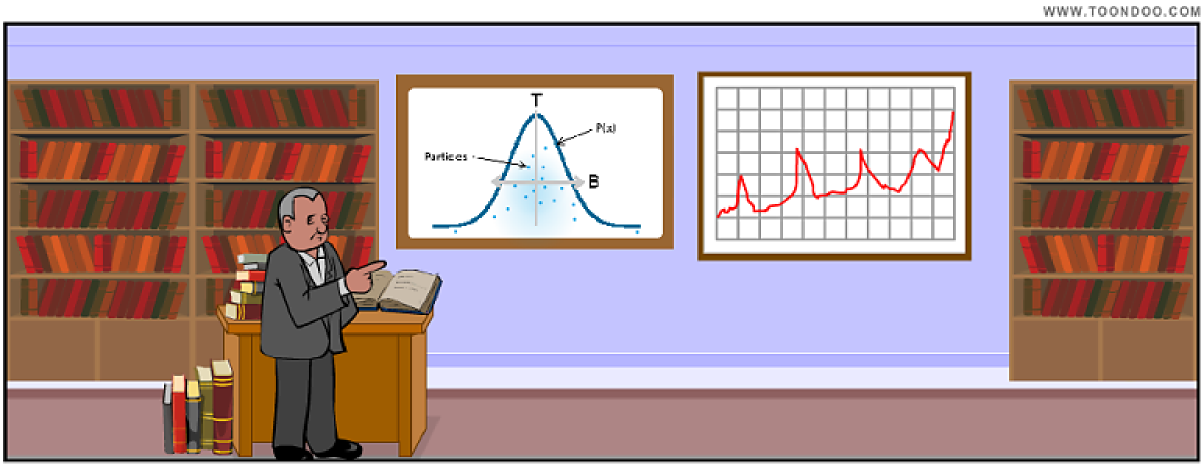
Basic Particle Filtering - Bayes Theorem

Dr Basic was downhearted when he realised that the observations would not be able to give the people of Assimilaris a good gift forecast. He considered resigning.

Whilst talking it over with a colleague one day, however, Dr Basic realised that the observations he was still painstakingly taking could be of use. In fact, he realised they could still be the key to solving the whole problem of the gift forecasts.

Dr Basic thought about Lord Ensemble's approach of using the King's equation to find **p(x)** and saw that he could still do this but - crucially - he could find the probability distribution of the particles' values *given* the observations. That is, he could find **p(x|y)**.

This would give him a much better understanding of what was going on up on the mountaintop.



Once he had decided to focus his efforts on finding **p(x|y)**, Dr Basic thought immediately that the logical place to start might be with Bayes' Theorem. Dr Basic quickly dug out an old textbook and found the expression. Substituting in the observations **y** and the particles **x** gave him:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \qquad (4)$$

Basic Particle Filtering - Particle Locations

Encouraged by his first step Dr Basic turned his attention to the individual terms of this new equation to see if there was anything that he could solve.

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$

He was immediately struck by the lurking presence of **p(x)** - that exact same probability distribution that had caused Lord Ensemble all of his trouble. What came next was a stroke of genius that Dr Basic didn't appreciate at the time but which quickly became central to his particular way of doing things.

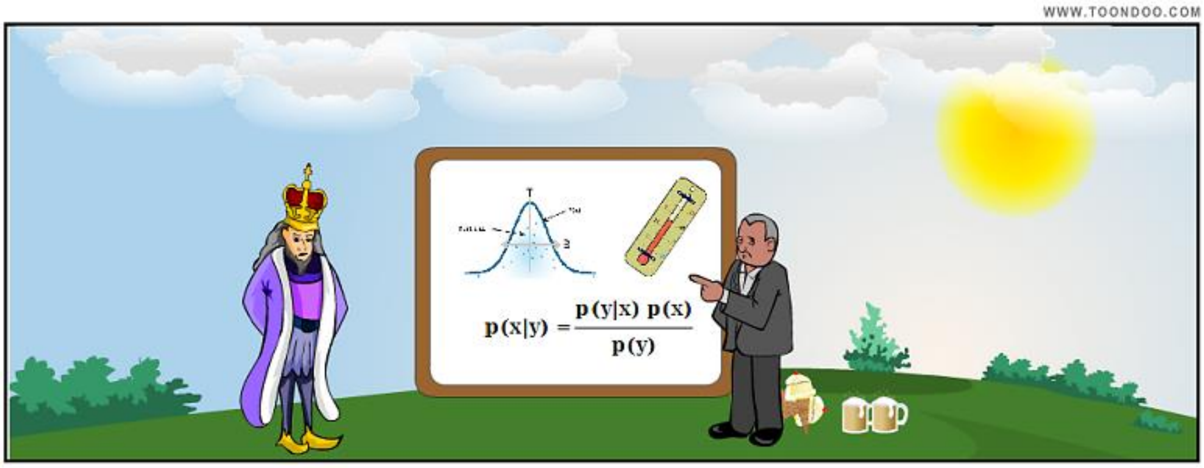
Unwilling to address Lord Ensemble's problem of finding the right kind of distribution to fit **p(x)**, Dr Basic, decided to simply let the positions of the particles define their distribution ("at least for now", he told himself at the time). In practise, Dr Basic defined **p(x)** as the sum of Delta Functions at each particle:

$$p(x) = \frac{1}{N}\sum_{i=1}^{N}\delta(x - x_i)$$

Dr Basic was very happy with his new breakthroughs and rushed to the King to tell him of his progress so far.

Basic Particle Filtering – Weights

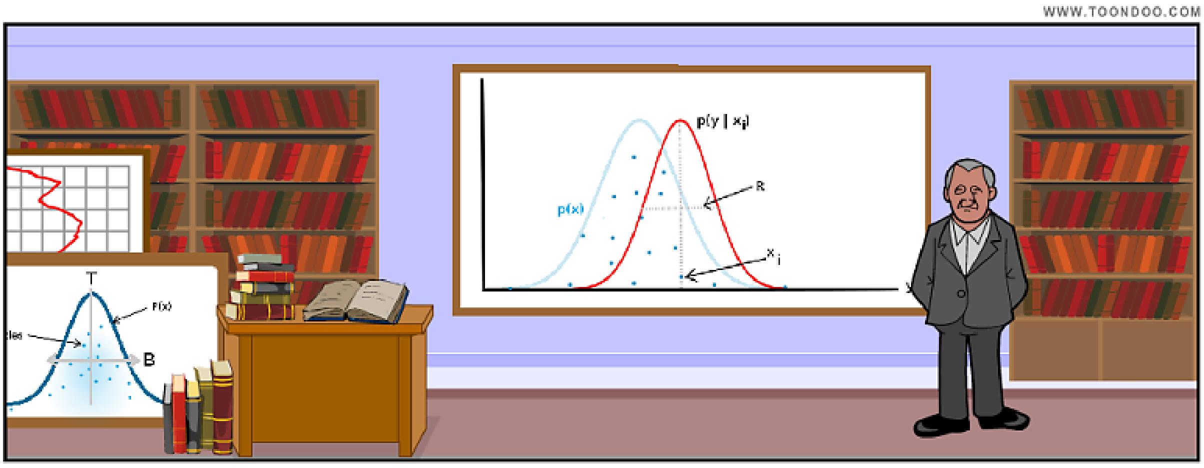
The King was very intrigued by Dr Basic's progress and asked him many questions about the last two equations. The King was particularly interested to know why there were two other terms in the expression for **p(x|y)** in equation 4. What other things - aside from the location of the particles - could possibly influence the distribution of **p(x|y)**?



Dr Basic was unable to answer the King's burning question and so returned back to his office to ponder the two remaining terms of equation 4.

After a few days of looking over the first two terms in the equation, Dr Basic realised that **p(y | x)** was the likelihood that the observations have the probability distribution of **p(y)** given the position of each individual particle. After some research in the library and old textbooks Dr Basic found that for a given particle i, this was given by:

$$p(y|x_i) = Aexp\left(-{}^{1}\!/_{2}\frac{(y-x_i)^2}{R}\right) \qquad (6)$$



Which left only **p(y)** to be found. If Dr Basic had known exactly what the true temperature on the mountaintop was then this would have been relatively easy to calculate, but as it was he was forced to take another approach.

Dr Basic reasoned that if he knew the probability of every single possible value of **x** and the likelihood of a given **p(y)** given all those values, then he could reasonably combine these to work out what the real **p(y)** actually was:

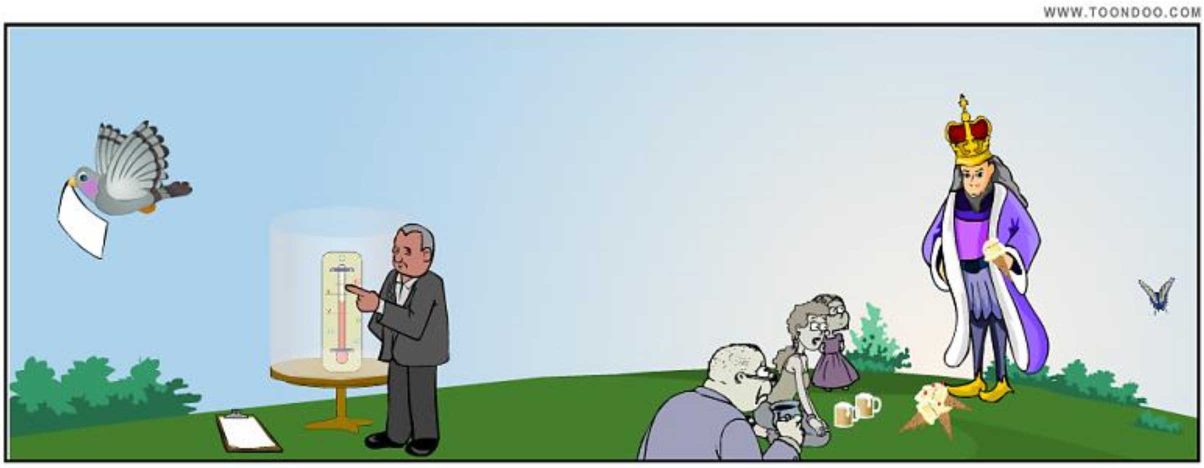
$$p(y) = \int p(y|x)p(x)dx \qquad (7)$$

And because **p(x)** has already been defined by the delta function in equation 5 - which integrates to unity - this simplifies down to:

$$p(y) = \frac{1}{N}\sum_{i=1}^{N} p(y|x_i) \qquad (8)$$

And substituting into this from equation 6 gives:

$$p(y) = \frac{1}{N}\sum_{i=1}^{N} A\, exp(-{}^{1}/_{2}\frac{(y-x_i)^2}{R}) \qquad (9)$$



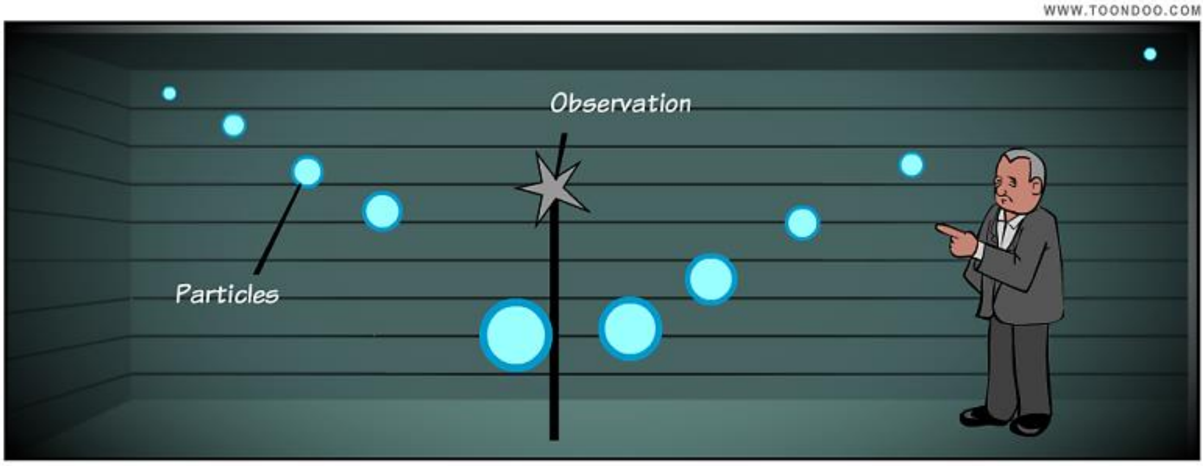
Putting these two expressions for **p(y | x**$_i$**)** and **p(y)** back together and cancelling the constant **A** gave Dr Basic the following equation:

$$\frac{p(y|x_i)}{p(y)} = \frac{\exp(-{}^{1}/_{2}\frac{(y-x_i)^2}{R}}{\frac{1}{N}\sum_{i=1}^{N} exp(-{}^{1}/_{2}\frac{(y-x_i)^2}{R})} \qquad (10)$$

Dr Basic looked at the terms and saw that the top half of the expression would be large if the values of **x**$_i$ and **y** were similar. He also saw that the bottom half would act to normalise this value for each particle in such a way that the sum of **p(y | x**$_i$**)/p(y)** for all particles at a given timestep would be unity.

And so this was how Dr Basic made his second great discovery. He saw that - other than the absolute positions of the particles - the additional information that was required to find **p(x|y)** was the relative proximity of the particles to the observation and a measure of the error in the observation (**R**). A particle lying very close to an observation with little error would have the greatest impact on the distribution.

As he was the son of a greengrocer, Dr Basic decided to call this particular feature of a particle it's "weight". The closer a particle was to the observation in comparison to the other particles, the greater it's weight. If all particles were equidistant from the observation they would all have equal weight.



With this, Dr Basic created the fundamental particle filters equation:

$$p(x|y) = particle\ weights \times particle\ positions \qquad (11)$$

Or, as it more commonly came to be known:

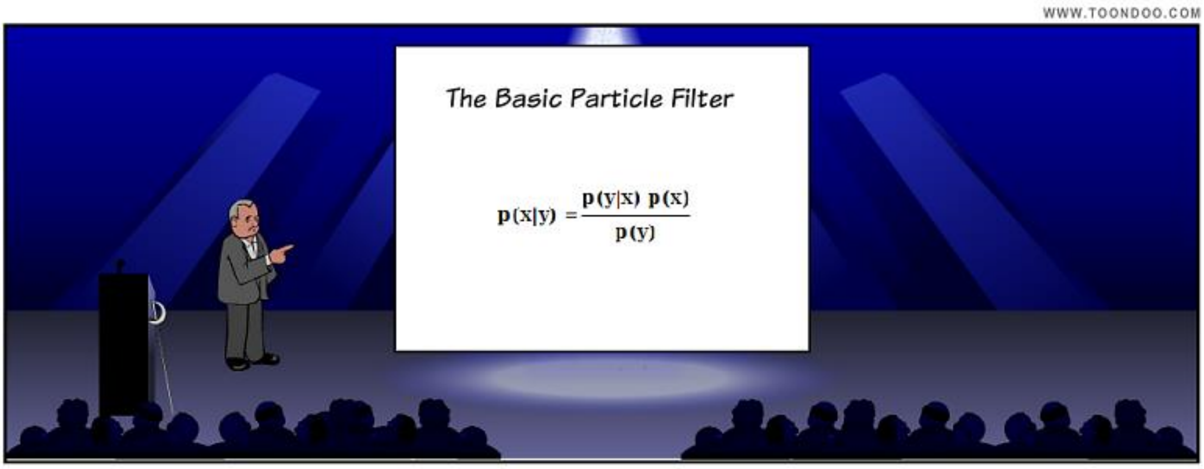
$$p(x|y) = \frac{1}{N}\sum_{i=1}^{N} w_i \delta(x - x_i) \qquad (12)$$

Where the weights are updated at each observation timestep by:

$$w_i = \frac{p(y|x_i)}{p(y)} = \frac{\exp(-{}^1/_2 \frac{(y-x_i)^2}{R}}{\frac{1}{N}\sum_{i=1}^{N} exp(-{}^1/_2 \frac{(y-x_i)^2}{R})} \qquad (13)$$

Basic Particle Filters in Action

And so it was that Dr Basic invented particle filters. The King was so grateful to Dr Basic that he gave him a knighthood. To celebrate the occasion, Sir Basic - as he was now known - demonstrated his basic particle filter in action.



Dr Basic showed how he had taken some of his temperature observations and compared them to the particles for 100 timesteps. At each observation he calculated the weights of the particles and by knowing these he could see which particles should be relied upon most heavily to predict the temperature up until the next observation.

He plotted these together, the truth, the observations and the particles to show the Basic Particle Filter in action. (Download Basic PF code to re-create the plots.)

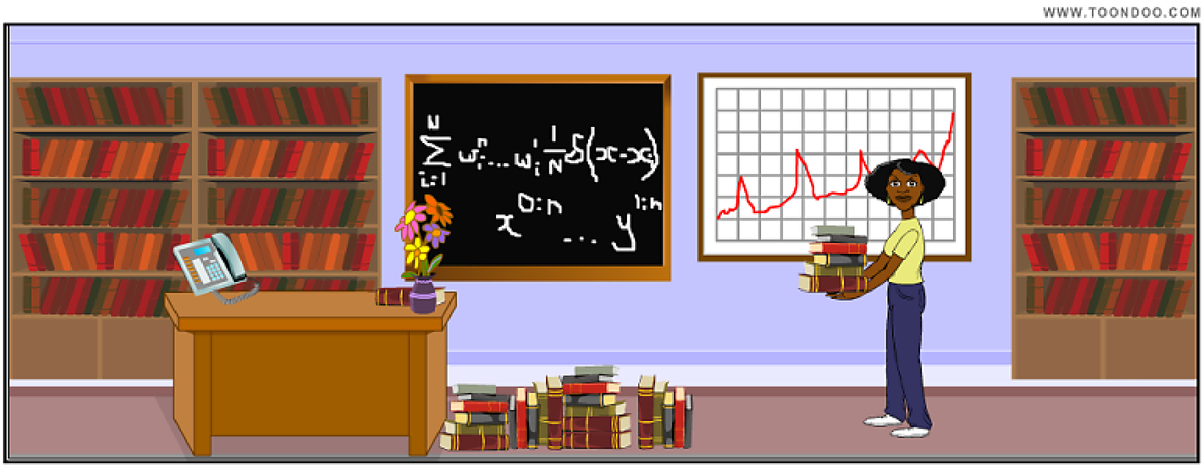
In his plots Sir Basic coloured the particles according to their weight, the more heavily they were weighted, the deeper their colour, this showed where the most representative particles were and didn't clutter the view.

As he brought up his last slide, the crowd erupted into a thunderous applause which lasted for many minutes. Clearly the basic particle filter was a success. The King and his kingdom were delighted and Sir Basic became a very popular man.

The Problem with the Basic Particle Filter

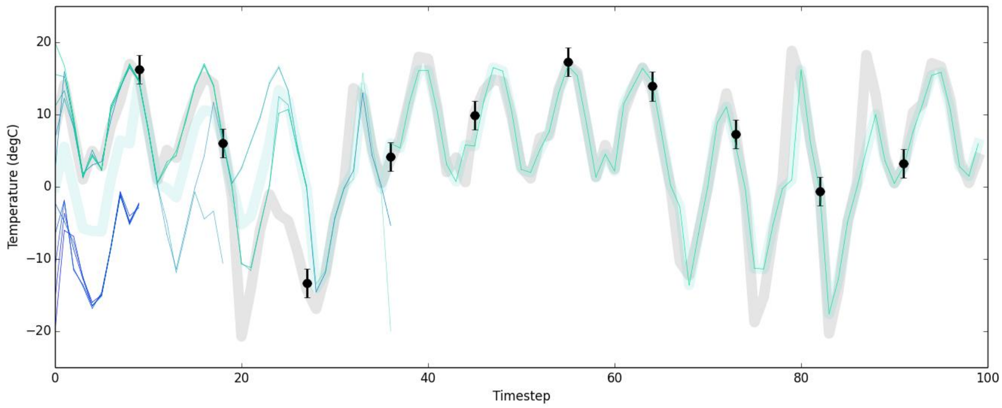
Sir Basic was at the top of his game and the Basic Particle Filter was working very well indeed. Wastage had been reduced and - though the forecasts were never entirely perfect, they were significantly better than before.

As he was given more and more work to do forecasting various new and important things using the Basic Particle Filter, Sir Basic took on an apprentice called Ms Standard.



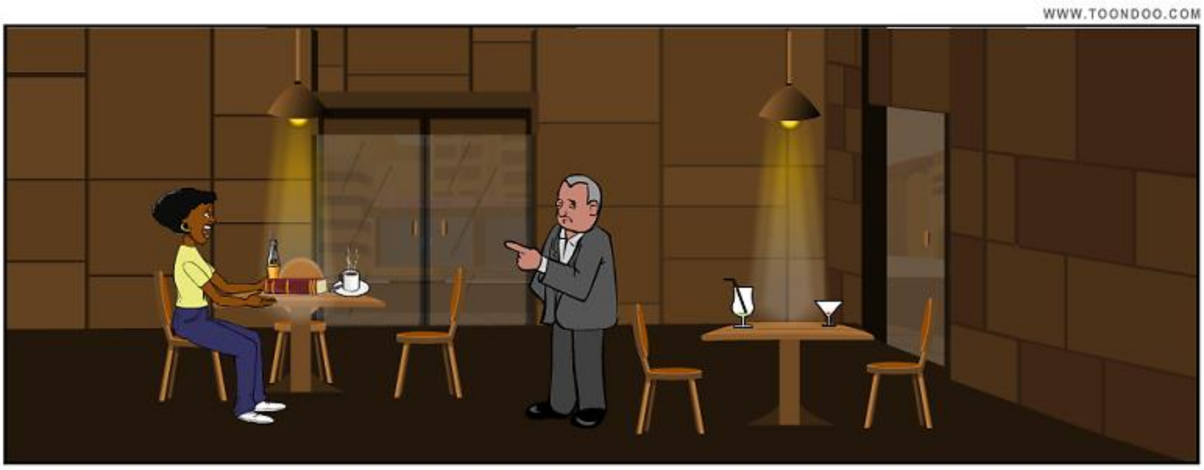
Ms Standard's main duty was running the models which produced the particles and using the Basic Particle Filter equations to produce the new forecast each timestep.

Ms Standard was very shrewd and quickly realised something that no-one else had noticed. Looking again and again at the timeseries of particles, Ms Standard noticed that only a few timesteps after initialisation, most of particles were given so little weight as to be completely negligible. Most of her day was spent calculating the values of particles which were actually never used in the forecast at all.



This is a major problem with the Basic Particle Filter, a problem which ultimately renders the extra particles themselves useless and any information regarding the probability distribution of the truth is totally lost as you end up with a single value.

Ms Standard went to see Sir Basic and told him what she had found. He took the bad news well and - after examining the evidence - agreed that this was indeed a problem.



However his years of hard work in the field and enormous popularity made him reluctant to tackle this new problem himself. He explained that this was a valuable opportunity for Ms Standard to advance her career and once she had solved it, he would step down and allow her to work in his place.

The Standard Particle Filter

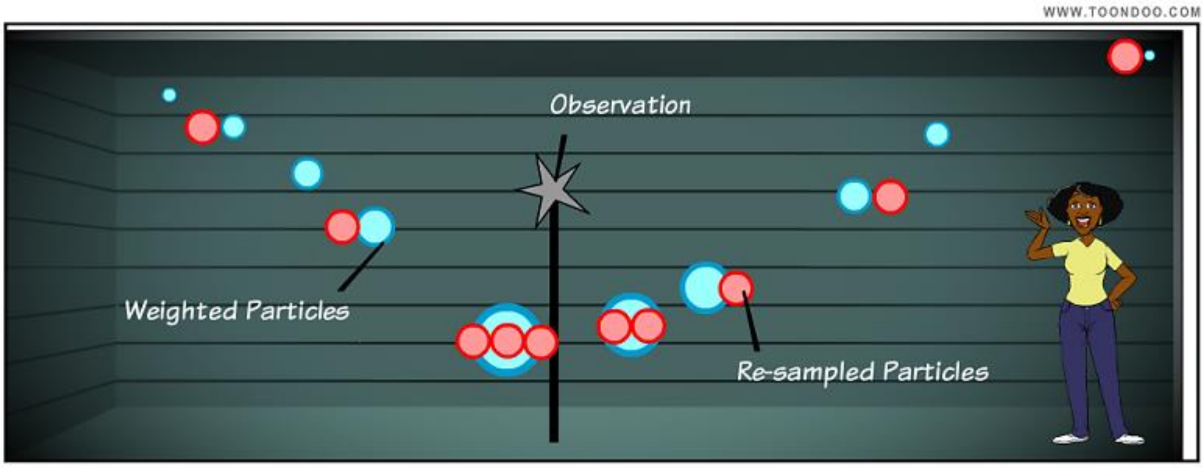
Ms Standard immediately set to work on improving the Basic Particle Filter. She soon realised that it was the information carried by the particles that was the key, rather than the particles themselves. Therefore - so long as the information was retained - there was no reason why individual particles couldn't be changed during the process.

And so it was that Ms Standard hit on the idea of re-sampling. At every observation, she analysed the weights of all the particles and calculated **p(x|y)** using Dr Basic's equation:

$$p(x|y) = particle\ weights \times particle\ positions$$

Then, she created a new distribution of particles, all with equal weights, chosen from the original distribution. Ms Standard used the method of Kitagawa (1996) (though there are other possible methods).

Practically speaking, if a particular particle was very close to the observations and therefore had a heavy weight, then during re-sampling many new particles would probably be created with this same value. On the other hand, if a particle had almost no weight at all, then it may not be re-sampled at all.

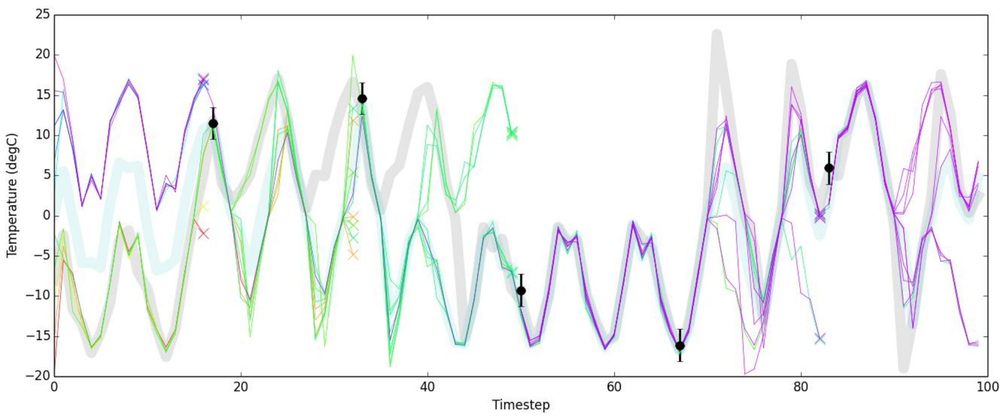


The new particles were each given equal weights and then taken forward in the next steps of the particle filter. At the next observation, the process would repeat again.

By doing this re-sampling, Ms Standard was effectively translating information held by particle's weights into information held by the particle's position. And she was getting rid of those irritating particles which contributed almost no information to **p(x|y)**.

Standard Particle Filters in Action

After checking her method to make sure that it was robust, Ms Standard went to see Sir Basic one morning and showed him her work.



(The crosses in the timeseries show where information from particular particles is discarded. The colours in both of these plots have been changed to clarify the change in particles over time)

She showed him that by performing re-sampling the distribution of particles was kept around the truth and that the process of re-sampling didn't affect the mathematics behind the particle filters at all.

(Re-create plots using Standard PF Code.)

Sir Basic was very impressed indeed. Cheerfully he declared over his cup of coffee,

"You have created the Standard Particle Filter. Who would have thought that it would happen so quickly?"

Ms Standard replied, "actually, I was thinking of naming it the Sequential Importance Resampling (SIR) Filter."

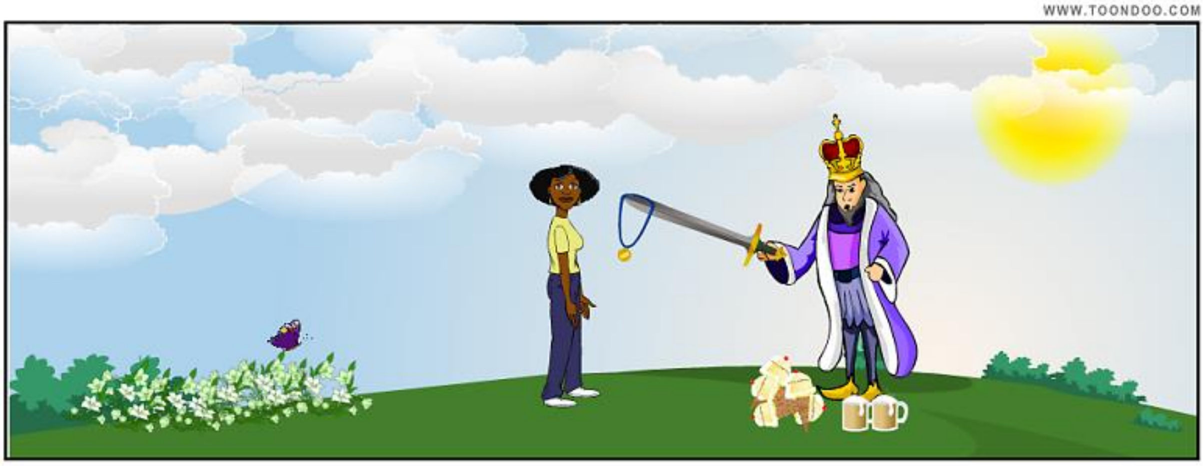
"Have it your way," said Sir Basic.

Sir Basic went on to explain that he would keep his promise and made arrangements to visit the King that afternoon. Ms Standard was delighted to be recognised for her work and excited to be in such an eminent position after only a short time at the Royal Department for Gifts. Sir Basic was equally delighted to find someone he could trust to run the department, allowing him to retire to the Bahamas.

The Problem with the Standard Particle Filter

Though she did not seek the same publicity as Sir Basic, Ms Standard's work was celebrated as visionary. Not least because of the huge cost benefit savings of making the most out of each individual particle between observations.

The King was very pleased with her work and awarded her a damehood.



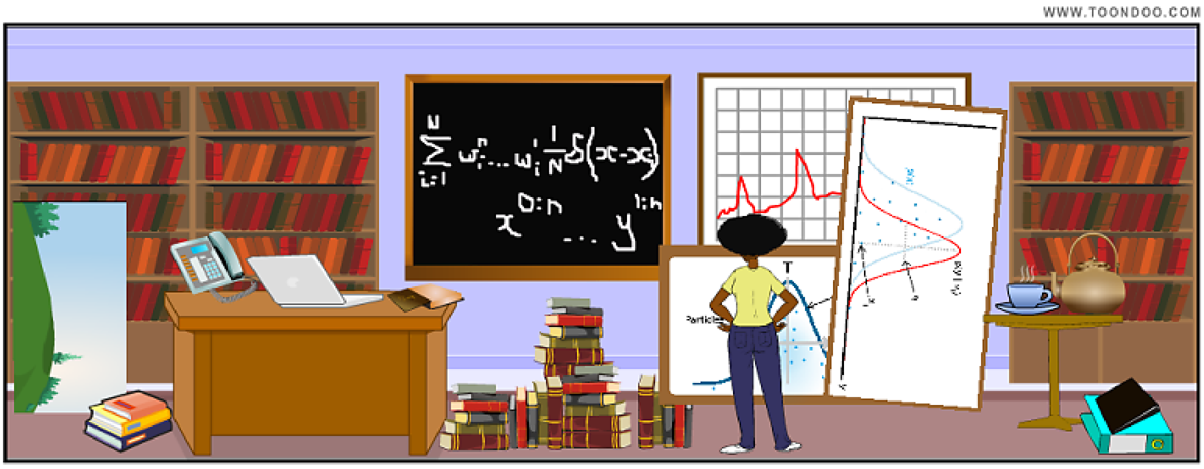
When she went to visit him on the mountain to receive her damehood, the King told her that - since it had been many years ago that he had calculated the equations for temperature - he had been calculating the equations for other phenomena.

The king now had expressions for over fifty more phenomena, including precipitation, wind direction, grass colour and bee population to name but a few. He believed that all of these things affected his well-being and his subsequent desire for hot chocolate or ice cream. The king benevolently insisted that these should also be included in the Standard Particle Filter.

Dame Standard was quite overwhelmed by all these additional variables. She knew it would take years to incorporate all of these into the King's Equation (or model) and the Department of Royal Gifts would need to take on a lot more staff.

Degeneracy

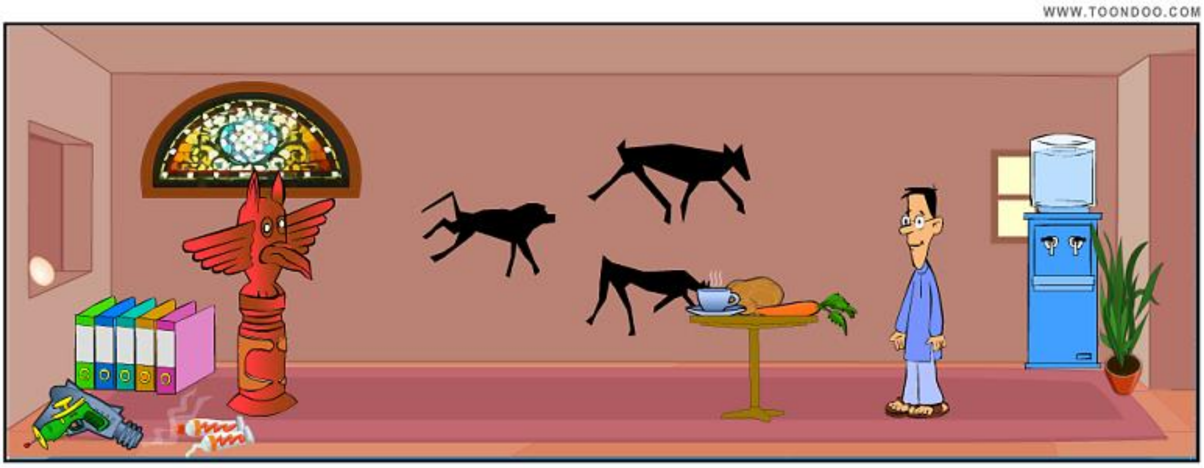
But most important to Dame Standard was that the particle filter could continue to incorporate observations into the King's Equations now that they were calculating changes in many many variables. After some initial tests she quickly realised that the Standard Particle Filter would no longer be up to the job with all of the new information that they needed to process. Eventually the same degeneracy that had blighted the Basic Particle Filter would affect the Standard Particle Filter and all new particles would ultimately be created from a single particle at every timestep.



Even though there were a few years to go, Dame Standard knew that she had to solve this problem as soon as possible to be ready for the new model. She also knew that she would be too busy herself to do this work, since Sir Basic had retired as he'd promised.

It was late one Friday night when Dame Standard made a bold decision. She decided to hire an Avant Garde, contemporary Data Assimilator. Known for his challenging techniques and bold new ideas...

Nudging the Particle Filter

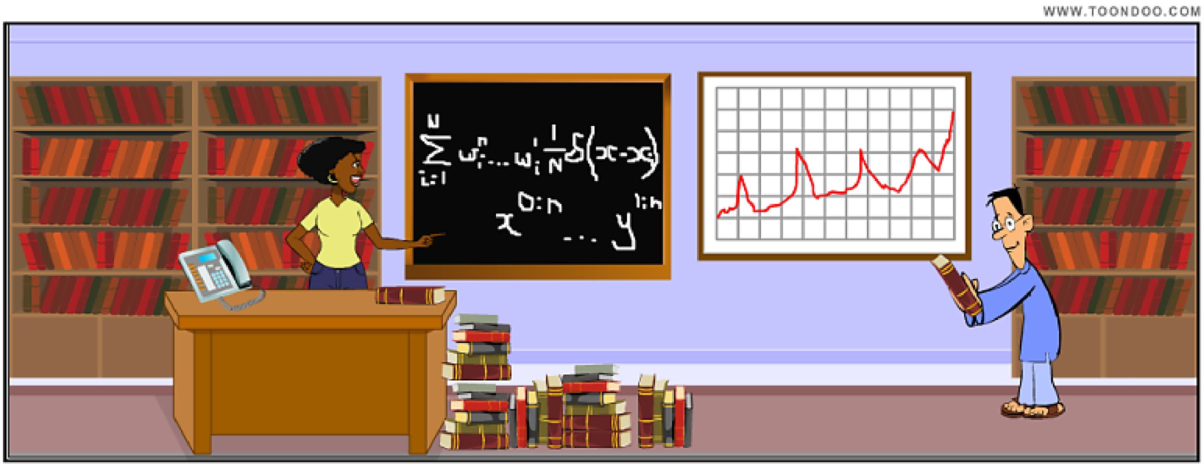


Dame Standard employed Nudge. He had chosen to forgo other names in order to be true to his craft. Truth be told, he was difficult to work with at times. But he was given a space to do his work and asked to come up with solutions to the Particle Filter degeneracy seen at multiple dimensions.

After only a few weeks at the Department for Royal Gifts, Nudge came to Dame Standard with a brilliant but incredibly controversial idea.

Nudge's thinking was that weighing the particles was simply not enough on its own. He suggested that a term be added to the King's equation for mountaintop temperature which would mean that each individual particle would move closer towards the temperature found by the observation. He suggested nudging the results to be more like the observations.

By nudging the results like this, he argued, there would be a stronger cluster of particles around the observation when it came to re-sampling and so less information would be discarded at each timestep.



Dame Standard was furious. What was the point of the particle filtering, if all you were going to do was "fix" the results so that they matched the observations? How was "making up" the physics of the King's equation like this any better than the Standard particle filter? In a rage she ordered him out of the room to try a new idea.

But Nudge stood his ground. He explained that the additional nudge term would be accounted for in the weighting. The more strongly a particle was nudged, the lower its weight would be. Nudge explained that he wasn't "making up" any data, all the same information would be retained. This method meant that information about the particle's trajectory was transferred to the particle's weight while the value of the particle (i.e. temperature it was representing) would be in the much better place relative to the observation.

Re-assured by this, Dame Standard asked to see his workings:

Nudge explained that he was not adding anything or fixing anything in the particle filter. He began his explanation with the fundamental particle filter equation:

$$p(x|y) = particle\ weights \times particle\ positions$$

and it's more popular derivative:

$$p(x|y) = \frac{1}{N}\sum_{i=1}^{N} w_i \delta(x - x_i)$$

He reminded Ms Standard that this equation was derived from the result of Bayes' Theorem:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$

Nudge explained that he had originally been interested in the **p(x)** term. He saw that the whole timeseries could be treated like a Markov Chain where each timestep is only dependent on the timestep before. So **p(x)** could be described as the probability density function of the whole chain **x** = {$\mathbf{x^0}$, $\mathbf{x^1}$, ... $\mathbf{x^{n-1}}$, $\mathbf{x^n}$}. This could be broken down into the probability of moving from state $\mathbf{x^0}$ at time 0 to the state $\mathbf{x^1}$ at time 1 - $\mathbf{p(x^1 \mid x^0)}$ - followed by the probability of moving from state $\mathbf{x^1}$ to state $\mathbf{x^2}$ etc. So that:

$$p(x) = p(x^n|x^{n-1})p(x^{n-1}|x^{n-2}) \ldots p(x^1|x^0)p(x^0) \qquad (14)$$

Because of this property, each of these individual terms could be calculated by the King's equation:

$$\boldsymbol{x^{n+1} = 0.5x^n + 25\left(\frac{x^n}{1 + (x^n)^2}\right) + 8cos(0.8n) + \beta^n}$$

Or as Nudge preferred to write it:

$$x^n = f\left(x_i^{n-1}\right) + \beta^n \qquad (15)$$

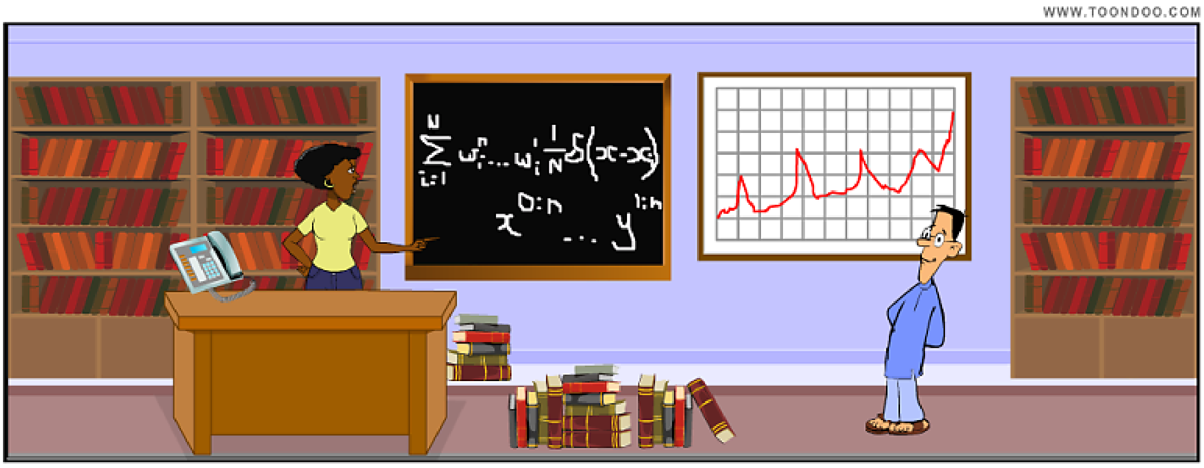
and he found through analysis that the variance of the random error $\boldsymbol{\beta}^n$ was equal to **Q**.

This expression showed Nudge that the probability distribution of **x** given the position of a particular particle at the previous timestep - $\mathbf{x}_i^{n-1}$ - could be given by:

$$p\left(x^n \middle| x_i^{n-1}\right) = N(f(x_i^{n-1}),Q) \qquad (16)$$

and therefore the probability of picking the current value of $\mathbf{x}_i^n$ from that distribution is given by:

$$p\left(x_i^n \middle| x_i^{n-1}\right) = A\ \exp(-{}^1/_2 \frac{(x_i^n - f(x_i^{n-1}))^2}{Q} \qquad (17)$$



Nudge glanced up at Dame Standard, she was nonplussed, having known all of this for quite some time.

"Then," said Nudge, "I had a vivid and compelling dream about conical valleys and black holes. I woke up wondering how these probabilities would change if I gave each particle a small pull towards the observations."

Nudge explained that with this pull the equation for $\mathbf{x}_i$ would change to:

$$x_i^n = f\left(x_i^{n-1}\right) + pull_i + \beta_i^n \qquad (18)$$

And that you would therefore have a new probability distribution for $\mathbf{x}^n$ which must include the fact that by nudging you must already know something about the observations. "I call it the 'proposal density'," said Nudge. "Or 'q' for short."

$$q\left(x^n \middle| x_i^{n-1}, y\right) = N(f\left(x_i^{n-1}\right) + pull_i, Q \qquad (19)$$

So the probability of picking the current value of $\mathbf{x}_i^n$ from this proposal density distribution is given by:

$$q(x_i{}^n | x_i^{n-1}, y) = A\ \exp(-\frac{1}{2}\, \frac{[x^n{}_i - (f(x^{n-1}{}_i) + pull_i]^2}{Q}) \qquad (20)$$

"So I tried it, to see what would happen," said Nudge. "I took Sir Basic's original equation.."

$p(x|y) = particle\ weights \times particle\ positions$

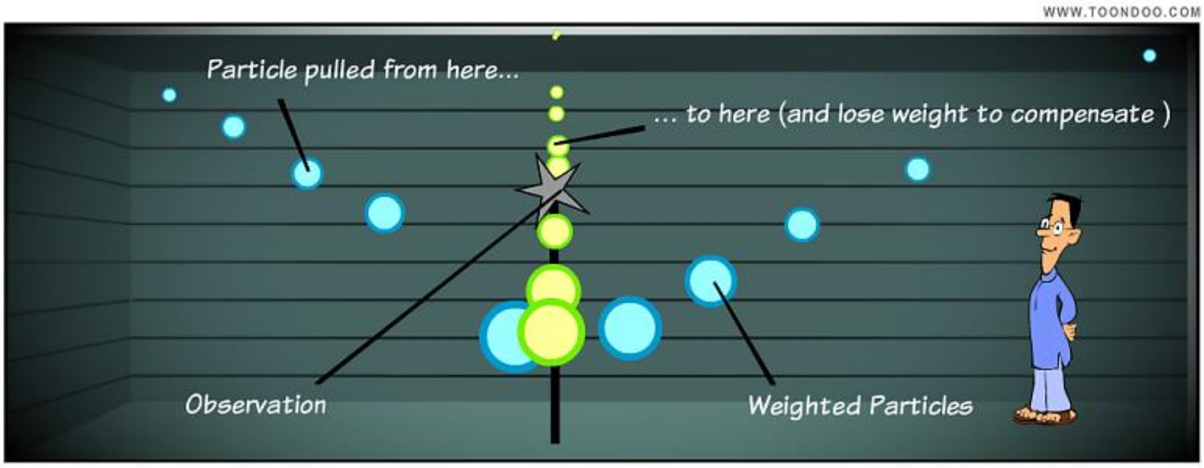
... and just set the following to be true for each particle:

$$position\ of\ particle\ i = q(x_i^n|x_i^{n-1}, y)\ q(x_i^{n-1}|x_i^{n-2}, y) \cdots q(x_i^1|x_i^0, y)p(x_i^0)$$

$$= \prod_{n=1}^{end\ of\ run} q(x_i^n|x_i^{n-1}, y)p(x_i^0) \qquad (21)$$

Seeing the thunderous look on Dame Standards face, he then quickly said "but in order to keep the original equation true I changed the weights accordingly - by multiplying with $\boldsymbol{p(xn|xn-1)}$ and dividing by $(\boldsymbol{xn|xn-1, y})$. You see? Using equation 13 as my base, the value of w_i then becomes:"

$$w_i = \frac{p(y|x_i)}{p(y^n)} \prod_{n=1}^{end\ of\ run} \frac{p(x_i^n|x_i^{n-1})}{q(x_i^n|x_i^{n-1}, y)} \qquad (22)$$

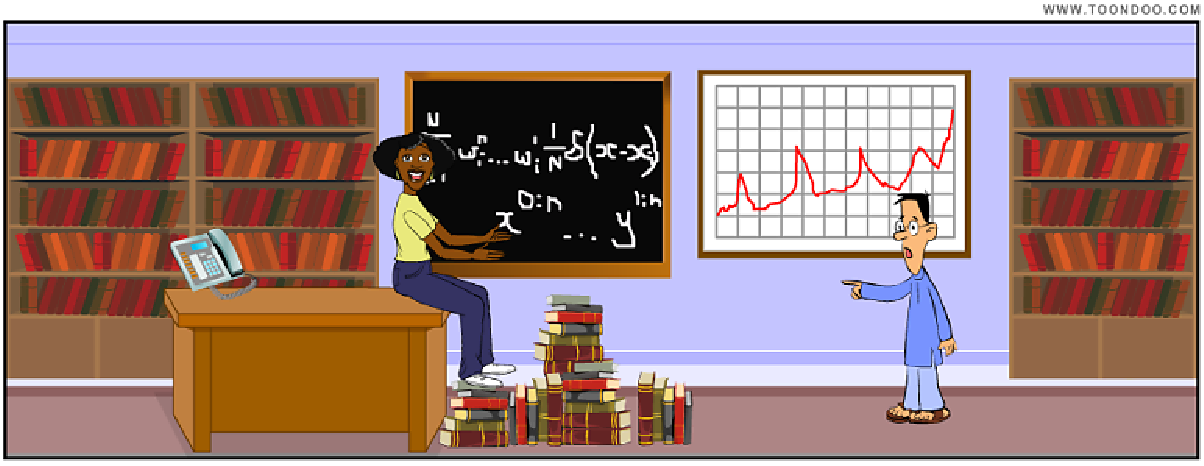


".. and so the full final equation becomes..."

$$p(x|y) = -\sum_N^1 {}_{i=1}^N \delta(X - X_i) \frac{p(y|x_i)}{p(y)} \prod_{n=1}^{end\ of\ run} \frac{p(x_i^n|x_i^{n-1})}{q(x_i^n|x_i^{n-1}, y)} \qquad (23)$$

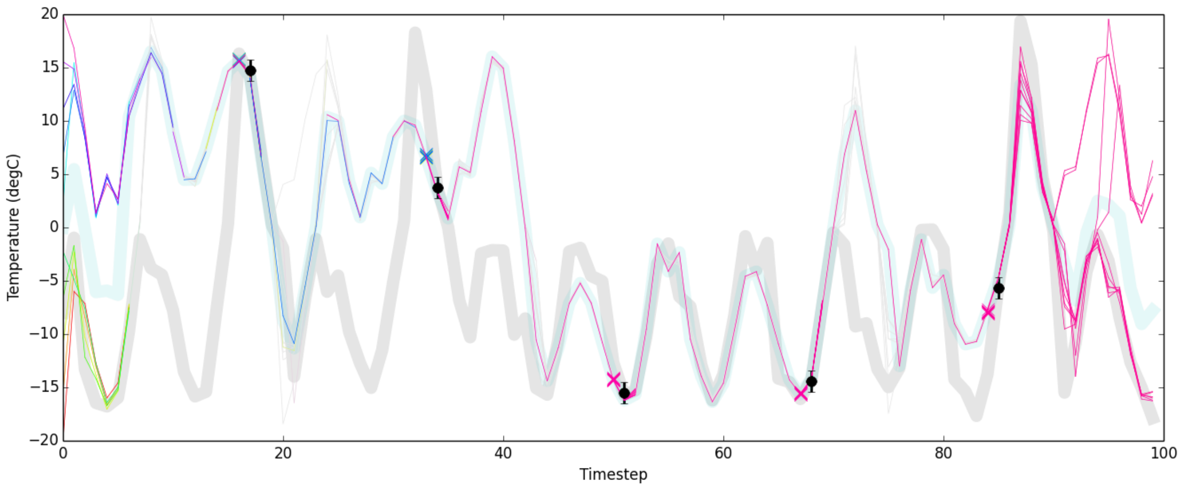
"And," beamed Nudge, "we can explicitly calculate this value at ever timestep using equation 6, equation 9, equation 19, equation 17 and equation 20."

Ms Standard agreed that Nudge's idea was indeed brilliant.

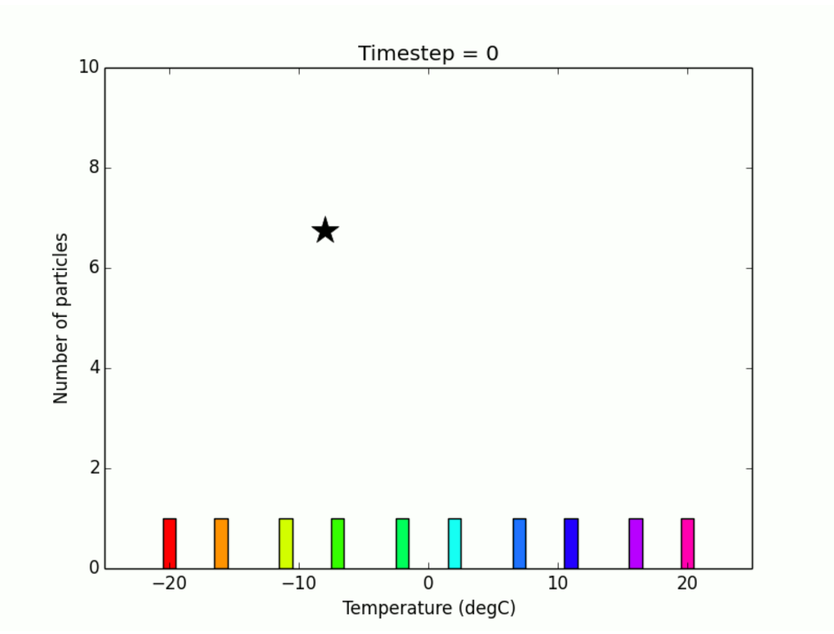


"But" she said - reluctant to burst his bubble - "if all you're doing is moving the information in each particle from the position to the weight, then surely you're going to get the same degeneracy? Now the particle won't be negligibly small because it's so far away from the observation, but more because it's weight is incredibly small. We're still pretty unlikely to sample it."

To back up her point, she asked to see Nudge's timeseries plot - where it was clear that the particles were ultimately being re-sampled from a small part of the initial distribution, effectively giving the same problem as experienced before:



The animated gif of particle histograms also showed the same thing from a different angle:

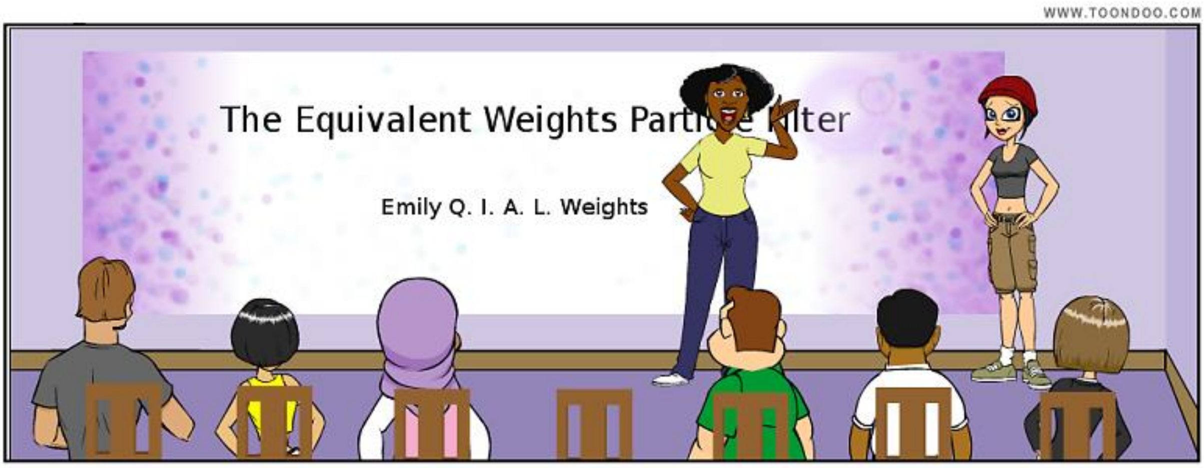


There was a long pause. Then Nudge said "good point". Without saying another word he picked up his things, nodded a thanks Dame Standard for the opportunity, and left.

Though she wasn't unhappy to see Nudge move on to pastures new, Ms Standard was still very pleased with his work - he had broken through into a new way of thinking about particles and the option of altering them into a more favourable position and then accounting for that alteration in the maths of the particle was a breakthrough.

Now that they had a new framework, Ms Standard was sure they could hit on a solution to make particle filters - which were already pretty good - the best possible tool for creating the King's climatology. But time was short and so she paid town criers to circulate an advertisement far and wide in search of the solution to the degeneracy problem.

After two weeks of frantic work, Emily Weights stood before Ms Standard and the other gift advisers to present her results.



After Dame Standard's introduction, an expectant hush passed over the advisers as they awaited her presentation. She began with Sir Basic's fundamental equation:

$$p(x|y) = particle\ weights \times particle\ positions$$

She reminded the advisers that over the past year the positions of the particles had come to be found by:

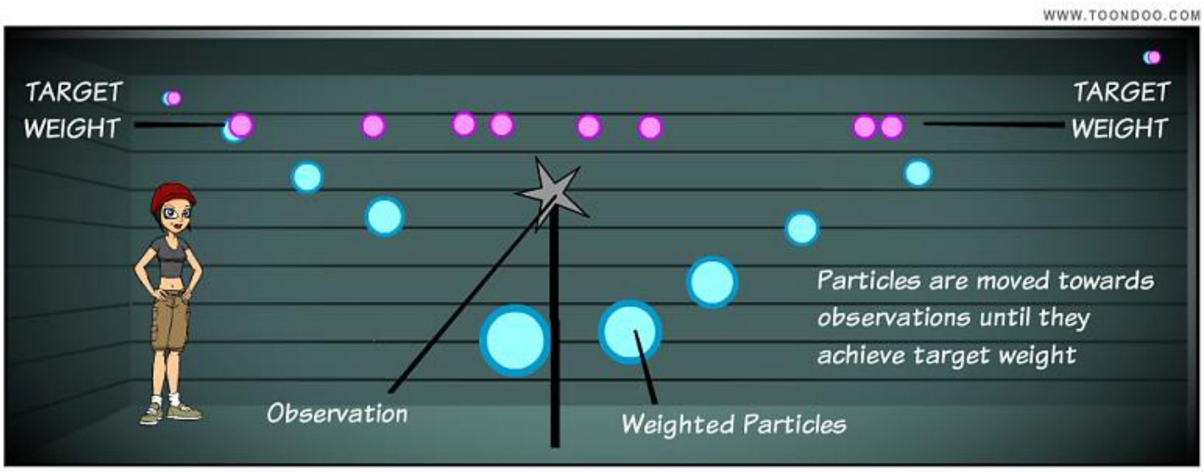
$$positions = \prod_{n=1}^{end\ of\ run} q\left(x_i^n \middle| x_i^{n-1}\right) p\left(x_i^0\right)$$

$$= \frac{1}{N} \sum_{i=1}^{N} \delta(x^n - x_i^n) \qquad (24)$$

And that the weights were given by:

$$weights = \frac{p(y|x_i)}{p(y)} \prod_{n=1}^{end\ of\ run} \frac{p\left(x_i^n \middle| x_i^{n-1}\right)}{q\left(x_i^n \middle| x_i^{n-1}, y\right)} \qquad (25)$$

Emily explained that is was her aim to move these particles in a way similar to Nudge's, but to find a position for each particle such that they almost all have the same weight. This would ensure that almost all particles are retained and re-sampled, reducing loss of particles.



Before embarking on this process Emily explained that - unlike the Standard Particle Filter - using Nudge's technique meant that each particles weights were changed between observations and the subsequent re-sampling. She made it clear that when a particle reaches the observation timestep then the starting weight of each particle could be called the "nudged weight", $\boldsymbol{w_i^{nudge}}$, and was defined by:

$$w_i^{nudge} = \prod_{n=1}^{(end\ of\ run)-1} \frac{p\left(x_i^n \middle| x_i^{n-1}\right)}{q\left(x_i^n \middle| x_i^{n-1}, y\right)} \qquad (26)$$

The first step in pulling each particle from it's individual nudged weight to a specified equivalent weight was simply choosing a weight which almost all particles could attain. As each particle could lose weight by being far from the observations and/or by being strongly pulled, there must be some point - between the particle's natural "un-pulled" trajectory and the observation - where is achieves maximum weight.

To find this weight, Emily first took the weight of each particle at the current timestep to be the following (she left out the normalising constant **p(y)** in order to keep things simple - it would cancel out eventually).

$$w_i^n = w_i^{nudge}\, p\left(y^{Obs}\middle|x_i^{Obs}\right) p\left(x_i^n\middle|x_i^{n-1}\right) \qquad (27)$$

Expanding this equation out using equation 6 and equation 17 gives:

$$w_i^n = \frac{w_i^{nudge}}{RQ} \exp\left(-\frac{1}{2}(y - x^n)^2\right) \exp(-\frac{1}{2}\left(x^n - f(x_i^{n-1})\right)^2) \qquad (28)$$

And taking the log of both sides:

$$\log(W_i^n) = \log\left(w_i^{nudge}\right) - \frac{(y-x^n)^2}{2R} - \frac{(x^n-f(x_i^{n-1}))^2}{2Q} \qquad (29)$$

This can then be reduced down to a quadratic equation in $\boldsymbol{x^n}$. Using the standard relationship to find the minimum in a quadratic equation, Emily found that the maximum possible of weight of any particle can be found by:

$$\log(w_{max}^n) = loh\left(w_i^{nudge}\right) - \frac{(y-f(x_i^{n-1})^2}{2(Q+R)} \qquad (30)$$

But the value of $\boldsymbol{w_{max}^n}$ would be different for every single particle. In order to retain a good number (say 80%) of the particles at each timestep, Emily had to choose a target weight that 80% of the particles could achieve.

Once she had identified the right target weight - $\boldsymbol{w_{target}^n}$- Emily then found the value of $\boldsymbol{x^n}$ which would give that weight. To do this, she used a well-respected solution published in van Leeuwen (2010). Using this published approach the value of $\boldsymbol{x^n}$ was given by:

$$x_i^n = f\left(x_i^{n-1}\right) + \propto_i K(y^n - f\left(x_i^{n-1}\right)) + \beta_i^n \qquad (31)$$

Emily explained that in this equation, the K was given by...

$$K = \frac{Q}{Q+R} \qquad (32)$$

... and that α was found by

$$\propto_i = 1 \pm \sqrt{1 - \frac{b_i}{a_i}} \qquad (33)$$

Where:

$$a_i = \frac{K(y-f(x_i^{n-1}))^2}{2R} \qquad (34)$$

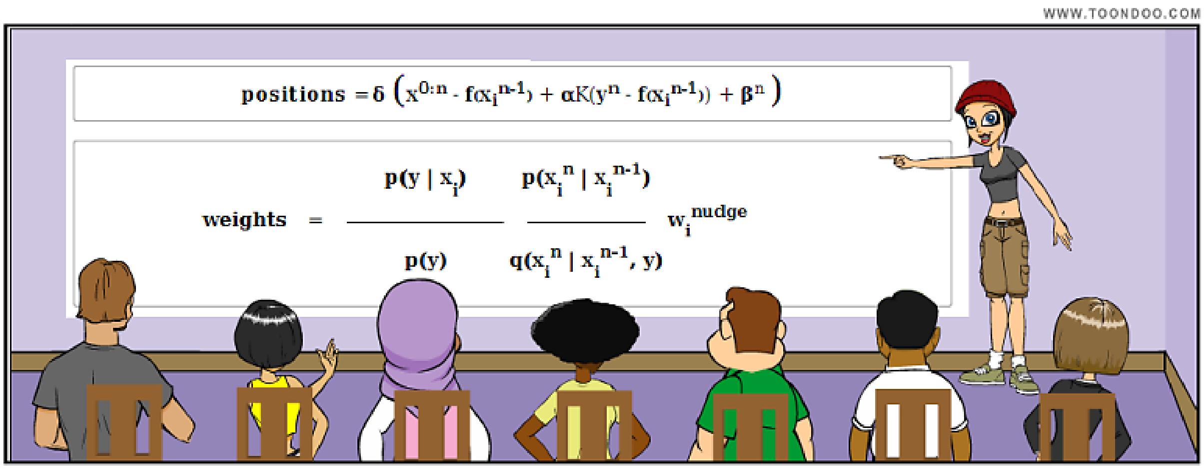
$$b_i = \frac{K(y-f(x_i^{n-1})^2}{2R} + \log\left(w_{target}^n\right) - \log(w_i^{nudge}) \qquad (35)$$

This, Emily explained, allowed her to calculate the amount of pull required on each particle to give it the target weight:

$$pull = \propto_i K(y^n - f(x_i^{n-1})) \qquad (36)$$

Finally, the act of pulling the particles towards the target weight introduced a small amount of error. The random error β in in the King's equation is responsible for this and means that following all these calculations the particles don't actually quite have the target weight. To account for this, the full weight must be re-calculated once at the end to find the true value for each particle using:

$$weights = \frac{p(y|x_i)}{p(y)} \frac{p(x_i^n|x_i^{n-1})}{q(x_i^n|x_i^{n-1}, y)} w_i^{nudge} \qquad (37)$$

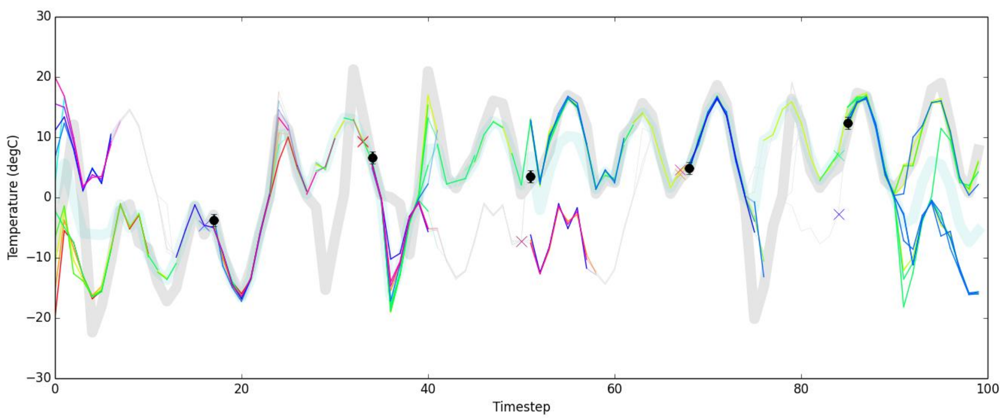


As Emily Weights spoke her last words a concentrated hush fell over the Gift Advisers. Emily was frozen to the spot.

"Does it work?" someone asked.

Emily quickly remembered her plots. "Oh yes," she said "I've tried it out with the king's equation and it works perfectly. It'll take some time to use it on the full multi-variable equations, but I'm confident that it'll be a huge improvement."

And with that she showed her plots of the changing particles with time...



... and a moving gif showing the histogram of particle weights and positions at each timestep:

The applause from the Gift Advisers was thunderous. They picked her up onto their shoulders and carried her off to the mountaintop where the delighted king immediately bestowed onto her a damehood and the golden shield of Most Valiant Knight of the Kingdom - the highest honour that he had ever awarded.

Emily was overwhelmed, Dame Standard was delighted and the day was declared a national holiday for ever more.

Many years of work still lay ahead of the Gift Advisers, but the Equivalent Weights particle filter promised to be exactly what they needed to perform data assimilation on the huge and complex non-linear environment that was their king's mountaintop.

"And that is how the Equivalent Weights Particle Filter came to be..."

Actually, the history of Particle Filters and the development of the Equivalent Weight Particle Filter is quite different! But our story should have taken you through the essential principles of the Equivalent Weights Particle Filter and it's predecessors. In reality the field of Particle Filters is a constantly evolving and very complicated subject.