Qualitative–Quantitative Reasoning: thinking informally about formal things

Alan Dix[1][0000−0002−5242−7693]

The Computational Foundry,
Swansea University, Wales, UK
`alan@hcibook.com`
https://alandix.com/academic/papers/ICTCS-QQ-2021

Abstract. Qualitative–quantitative reasoning is the way we think informally about formal or numerical phenomena. It is ubiquitous in scientific, professional and day-to-day life. Mathematicians have strong intuitions about whether a theorem is true well before a proof is found – intuition that also drives the direction of new proofs. Engineers use various approximations and can often tell where a structure will fail. In computation we deal with order of magnitude arguments in complexity theory and data science practitioners need to match problems to the appropriate neural architecture or statistical method. Even in the supermarket, we may have a pretty good idea of about how much things will cost before we get to the checkout. This paper will explore some of the different forms of QQ–reasoning through examples including the author's own experience numerically modelling agricultural sprays and formally modelling human–computer interactions. We will see that it is often the way in which formal and mathematical results become useful and also the importance for public understanding of key issues including Covid and climate change. Despite its clear importance, it is a topic that is left to professional experience, or sheer luck. In early school years pupils may learn estimation, but in later years this form of reasoning falls into the gap between arithmetic and formal mathematics despite being more important in adult life than either. The paper is partly an introduction to some of the general features of QQ-reasoning, and partly a 'call to arms' for academics and educators.

Keywords: informal reasoning · estimation · mathematical models · order of magnitude · covid models · monotonicity.

1 Motivation

When I first read Hardy and Wright's Number Theory [15] I was captivated. However, as much as the mathematics itself, one statement always stood out for me. In the very first chapter they list a number of "questions concerning primes", the first of which is whether there is a formula for the nth prime. Hardy and Wright explicitly say that this seems "unlikely" given the distribution of the series is "quite unlike what we should expect on any such hypothesis." I

think most number theorists would still agree with this assertion, indeed many cryptographic techniques would collapse if it such a formulae were discovered. Yet what is this sense that the structure of primes and the structure of formulae are so different? It is not formal mathematics itself, else it would be a proof.

In engineering, computation, physics, indeed any quantitative or formal domain, the precise and provable sits alongside an informal grasp of the nature of the domain. This was certainly true in my own early and more recent work on formal modelling of human computer interaction: sometimes, as in the case of undo, one can make exact and faithful statements and proofs, but more often in order to achieve formal precision, one resorts to simplistic representations of real-life. However, despite their gap from the true phenomena, these modes, however lacking in fidelity, still give us insight.

I'm sure this will be familiar to those working in other areas where theoretical models are applied to practical problems. There is a quantum-like tension between the complexity of the world and our ability to represent it, between accuracy and precision, between fidelity and formality. Yet, we do learn about real phenomena from these simplified models, and in many contexts, from primary school estimation to scientific research we use these forms of thinking – I call this qualitative–quantitative reasoning.

This has become particularly important during Covid, when both simple formulae and massive supercomputing models offer precise predictions of the impact of specific interventions. However, even the most complex model embodies simplifications and it is when the different models lead to qualitatively similar behaviours that they are most trusted. Similar issues arise for climate change, international economics and supermarket shopping.

Qualitative–quantitative reasoning is ubiquitous, but not often discussed – almost a dirty secret for the formalist and yet what makes theory practical. There are lessons for science and for schools, challenges for visualisation and argumentation. I don't know all of the answers, but by bringing this to the surface I know there are exciting questions.

In the rest of this paper we'll first move through a series of examples that each exhibit different forms of QQ-reasoning. The final section will outline practical and theoretical challenges.

2   Informal insights from formalism – the PIE model

My first work as an academic centred on creating formal models of interactive systems [9], notably the PIE model [8], a simple input–output models of interaction (see Figure 1). Whilst cognitive models try to model the mental behaviour of humans, the intention here was to model the systems that people use and to formalise key properties that lead to a system being usable.

Some aspects of this are amenable to strong proofs. Notably undo, which is expected to have predictable properties (really important it works!) but which also has a relatively straightforward algebraic definition:
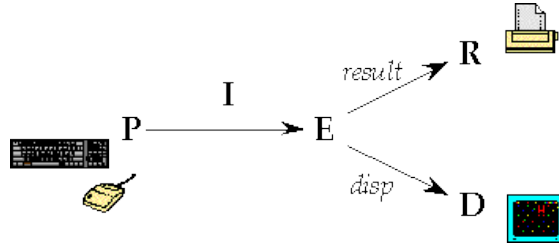
Fig. 1. The PIE model, a siple input–output modle of interactive systems

$$\forall c \in Commands : c \frown undo \sim null \tag{1}$$

(Here $\frown$ means performing the commands one after each other and $\sim$ means "has the same effect in all contexts".)

There are slight nuances to this, occasionally commands (such as typing) are clumped, and some purely presentation-level commands (such as scrolling) are ignored. However, this is pretty solid. One core question though is whether undo itself is undoable; that is:

$$undo \frown undo \sim null \tag{2}$$

Some systems *appear* to have this property, doing undo twice acts as a single-step redo. This is often called *flip undo*. However, one of the early proofs in the area showed that this was impossible for all but trivial systems [9]. To see this consider two commands c_1 and c_2:

$$
\begin{aligned}
c_1 &\sim c_1 \frown (undo \frown undo) \\
&= (c_1 \frown undo) \frown undo \\
&\sim null \frown undo \\
&\sim (c_2 \frown undo) \frown undo \\
&= c_2 \frown (undo \frown undo) \\
&\sim c_2
\end{aligned}
\tag{3}
$$

That is all commands have the same effect, which can only happen if the system has no more than two states.

As well as being a theoretical result it had the practical application that one should not attempt to 'debug' pure forms of flip undo to attempt to make undo just like any other command – this is impossible. Instead we have to accept that undo commands (undo, redo, etc.) have to be treated as a separate kind of command. In later work, Mancini's thesis used category theory to show that with fairly minimal assumptions, there are only two kinds of consistent undo

system: forms of flip undo (but where undo is treated as a special command) and stack-based undo–redo [6, 18]

However, these cases of complete proofs were comparatively rare. Many aspects of interaction are far more complex.

The initial impetus for the PIE model came from Harold Thimbleby's quest for 'Generative User interaction Principles' (GUEPs) [25] and also the desire for systems that were 'what you see is what you get'. This was used to formalise various forms of *predictability* and *reachability* properties, the former regarding whether it was possible to infer the state of the system, and the effect of commands from its display, and the later how easy it was to get to desired states (undo is related to this).
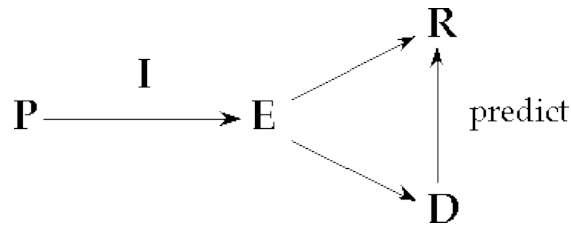
Fig. 2. A simple formulation of *predictability* in the PIE model – require a function from the display to the final result.

In relatively simple cases, such as medical devices, these properties can be verified by model checking [2], but this is impossible for larger systems. Even more critical issues such as the special undo commands become increasingly frequent: models that accurately model real systems rapidly become Baroque and those that are clean enough to reason with do not model reality.

This is not just a problem for interactive systems, but a general issue for modelling – complexity and simplicity at odds However, anyone who has created a formal specification of a substantial system will tell you that it is usually not so much the final specification that matters, but the understanding you gain through the *process*. Similarly. many theoretical treatments of issues are so far from the real system that they cannot in any way be used to predict precise behaviours, but nevertheless, the insights gained through theoretical analysis and proofs yield understanding that may help in more practical situations.

3   Making decisions – electrostatically charged agricultural crop sprays

Before modelling humans I modelled agricultural crop sprays.

Factory paint sprays often use electrostatically charged spray droplets, which are then attracted to an earthed object, such as a car, ensuring a full coating

and less waste. By a similar principle if an agricultural spray is charged it is attracted to the crop potentially leading to better coverage, less waste and less environmentally damaging spray drift. The main centres researching this in the early 1980s were ICI and the National Institute for Agricultural Engineering. At the latter we created numerical simulations of the movement of charged sprays in order to understand their behaviour and improve design choices [7].

Given computer speeds then were measured in KHz rather than GHz and memory in 10s of Kb, the models were, perforce, simple! However, even today the complexity of a field of swaying wheat would challenge a super computer. To make this tractable, the modelling was performed in two parts.
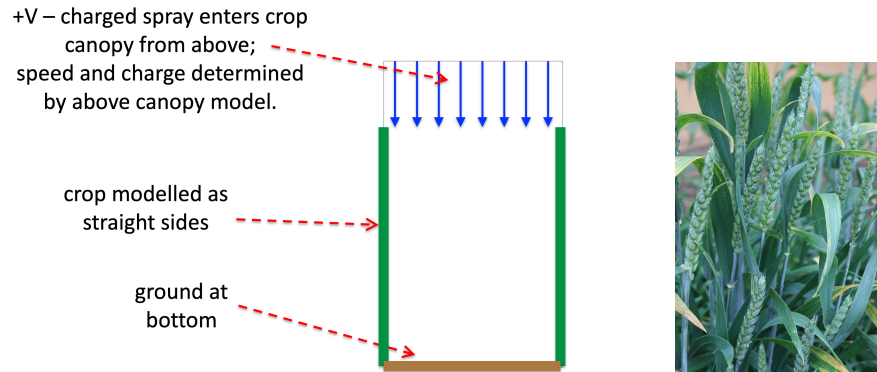
The first stage was to model the transport from crop sprayer to the top of the crop (Figure 3). Note that this has been flattened to two dimensions (effectively assuming a infinitely long spray boom!) to make the computation tractable. This is relatively simple, a point source for the spray, with an area held at high voltage (to represent the sprayer itself) and the crop top treated as a flat earthed surface, ignoring the fine structure. The output from this stage is the speed and density of the drops as they enter the top of the crop.
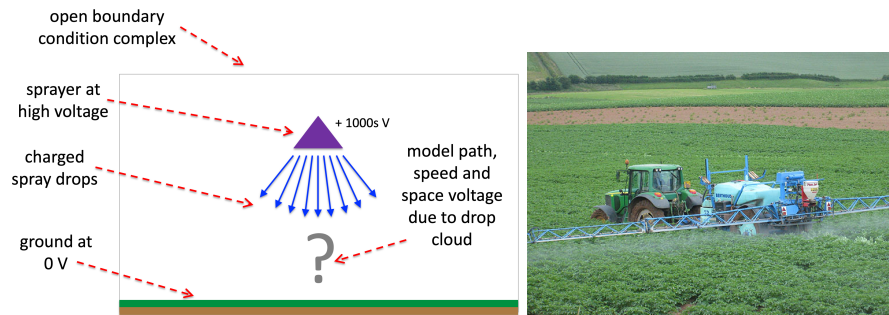
Fig. 3. Agricultural spray above the canopy (left) model of electrostatic spray (right) real sprayer (image right: Pauline Eccles (CC BY-SA 2.0); https://commons.wikimedia.org/wiki/File:Crop_spraying,_Bromsash_-_geograph.org.uk_-_1367703.jpg)

As is evident there are already several simplifications here. However, the within canopy modelling is far more difficult. In reality crops have leaves, seed heads and are dfferent sizes. In the model these are treated as flat vertical lines (Figure 4). Furthermore this as also a 2D model, so the crop is effectively modelled as infinitely long parallel metal plates. Indeed, for some experiments with real spray, such plates were used with paper collectors in order to obtain physical spray coverage data.

The speeds and density of drops entering the canopy from the above crop model could be used to match the relevant within-crop model in order to create an end-to-end model of how initial flow rates, drop size, charge etc. affect spray

+V – charged spray enters crop
canopy from above;
speed and charge determined
by above canopy model.

crop modelled as
straight sides

ground at
bottom

Fig. 4. Within the crop canopy (left) model of electrostatic spray (right) real wheat (image right: Stephencdickson (CC BY-SA 2.0); https://commons.wikimedia.org/wiki/File:Green_wheat.jpg)

deposition. The output data was copious, but was categorised into three classes (Figure 5):

Class I – great penetration: spray misses crop and mostly ends up in the earth

Class II – uniform deposit: spray creates a relatively even coverage

Class III – little penetration: spray ends up mostly at the top of the crop

Of these, it is (2) we want; both (1) and (3) effectively waste spray and may leave untreated potions of crop.
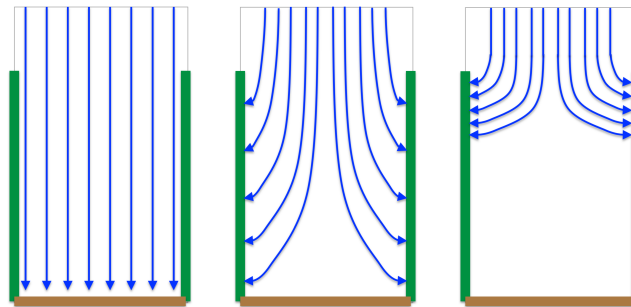
Fig. 5. Electrostatic spray coverage (left, Class I) misses crop and ends on the ground; (middle, Class II) good coverage of the crop; (right, Class III) too concentrated at the top of the crop.

In rough terms looking at the inputs to the *within canopy* model, the first class corresponds to fast or low charge droplets, and the last to slow or high charge particles, but this simple correspondance is more complex when looking at the complete above and within canopy models. High charge on small drops can lead to high space charge of 10s of thousands of volts (rather like the way a rain cloud builds up charge to create lightening) and this can then accelerate the drops as they enter the canopy so counter intuitively mean they end on the ground (Class I).

This knowledge itself was useful as it was hard to measure space charge. However, part of the aim was to go beyond the scientific knowledge to practical design advice. The mathematical model allowed one to make precise predictions as to which class a particular set of input parameters would yield, but of course the model was very far from reality. Instead dimensional analysis was used to reduce the input set to two main dimensionless features (π_1, π_2), the modelling runs were then plotted into the two dimensional design space and a map produced of how the input parameters corresponded to the classes, rather like the phase space of a gas–liquid–solid for water (Figure 6).
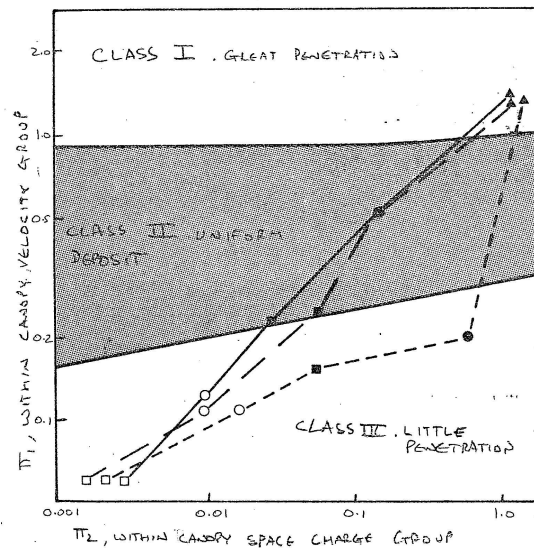

Fig. 6. 'Phase space' of different classes of spray pattern

While we had little confidence in the precise values of the modelling, the overall shape of this map was useful. For example, if we were getting too much spray on the ground (Class III), we might either try to increase dimensionless

parameter π_1 or to reduce parameter π_2, either of which could be manipulated using different concrete parameters.

Note how the very precise, but massively over simplified numerical model was used to create a qualitative understanding of the design space, which could then be used to make useful engineering interventions.

4   Orders of magnitude – climate change and complexity

4.1   Infinitesimals and limits

I recall reading Conway's "On numbers and games"[4] while still at school and being transported by the shear exuberance of the text. There can be a tendency to skip to the second half (part one) a joyous exploration of the odd aritmetic properties of games. However, the first half (part zero) is equally exciting dealing with, what has become known as, 'surreal numbers' – both transfinite ordinal arithmetic (fairly commonly taught in maths courses), but also (less commonly taught) the way this can also give a formal treatment of infinitesimals.

Even if you've not come across these formal infinitesimals, you will have been taught calculus using lots of ϵs and limit proofs:

$$f'(x) \; = \; \lim_{\epsilon \to 0}(f(x+\epsilon) - f(x)) \tag{4}$$

Crucially we learn that we can often ignore order of magnitude smaller terms: ϵ terms when dealing with 'ordinary' sized numbers, or ϵ^2 terms when dealing with ϵs:

$$f(x+\epsilon) - f(x) \; \approx \; \epsilon f'(x) \tag{5}$$

4.2   Day-to-day reasoning

In everyday life we also understand this, we may say "it's only a drop in the ocean". Formally we may use 'much greater than' ($\gg$) or 'much less than ($\ll$), but also informally we effectvely use rules such as as:

$$A \gg b \, \wedge \, b > c \;\; \implies \;\; A \gg c \tag{6}$$

and:

$$A \gg b \;\; \implies \;\; A + b \approx A \tag{7}$$

Unfortunately, less well understood in day-to-day logic is that the ocean is made up of drops, that is:

$$A \gg b_i \, \wedge \, N \, is \, very \, large \;\; \implies \;\; A \not\approx A + \sum_{i=1}^{N} b_i \tag{8}$$

In fact, there are 'thrifty' sayings that capture this: "many a mickle makes a muckle', or "mind the pennies and the pounds look after themselves". However, despite out best environmental or fair trade intentions, it is too easy when deciding on purchases in the supermarket, or choosing whether to walk or jump in car, to simply think "it won't make a difference". For ten thousand years, humanity was able to think like that, assuming that our individual impact would be absorbed by the vastness of land, sea and air. This underpins Locke's "as much and as good" proviso for the fair acquisition of land [17, Chap. V, para. 27], effectively assuming that nature's bounty is inexhaustible.

Of course, we now face the imminence of climate change, the ubiquity of plastics in the the oceans and, with Covid, the critical nature of thousands of personal precautions, each insignificant in themselves, yet between them allowing or preventing the spread of disease. Looking back, we are also able to see that these impacts, while ever-accelerating, are not entirely new; for example, it is possible that the desertification of central Australia was due to slash-and-burn farming by early settlers thousands of years ago [19].

4.3   Algorithmic complexity

In complexity theory, we argue formally about such order of magnitude relations using big and little 'O' notation. At a practical level we also get used to effectively counting the levels of directly or indirectly embedded loops to get an idea of the exponent r in $O(N^r)$.

Just like with plastic waste, we can sometimes forget that these are about theoretical limits and that in practice an $O(N^2)$ algorithm with a small contant K, may actually be faster than an $O(N \log N)$ with large K.

An extreme example of this is the linear programming simplex algorithm [5], one of the most successful early examples of operational research. Simple linear programming problems consist of N linear constraints over M variables ($N > M$). The optimal value of a linear objective function must lie at one of the vertices (Fig. 7). The simplex algorithm is basically a form of hill–climbing optimisation, moving from vertex to neighbouring vertex following the direction of maximum gain.

Given a linear objective function, the simplex algorithm is guaranteed to terminate after a finite number of steps, and in practice is linear in the number of constraints N. I say 'in practice', because in theory it can be much worse. Indeed it is possible to create Byzantine examples were the simplex algorithm visits all C_M^N vertices. That is its *worst case* behaviour is $O(N^{N-M})$.

In fact there are alternative algorithms for linear programming that have better worst case behaviour (I once heard of one that was $O(N \log N)$, but not been able to track it down). However, in practice they are all far slower in terms of *average case* complexity.
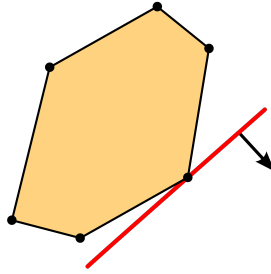
Fig. 7. Linear programming – the linear constraints create a polytope feasible region, the optimal value of a linear objective is on one of the vertices (image https://commons.wikimedia.org/wiki/File:Linear_optimization_in_a_2-dimensional_polytope.svg)

4.4   Sorting

Furthermore, the real world is finite. For some graph/network problems, where algorithms are often exponential or multiple-exponential, N more than five or six is enough to end up in the theory 'limits'. However, for other problems practical limits may be more significant.

We all know that sorting is $O(N \log N)$, but in fact every real sorting algorithm works on finite sized keys within a computer with finite disk space. When sorting finite keys, in principle bucket-sorts give algorithms with time linear in N. See, for example, the IBM Punch Card Sorter in action [23] – this required just W passes to sort W-character keys, that is effectively an $O(N W)$ algorithm.

You might wonder how this squares with the well-known information-theory-based $O(N \log N)$ lower bound for sorting algorithm. First, the theoretical bound depends on it being necessary to compare sufficient items to determine a *total order* on the items. If $W < \log N$ there will be many equally placed items. Second, the information theory bound is incredibly broad, even working with magical oracles that tell you where to put items – effectively it is lower bound on the time taken to *read the result*. Even with bucket sorts you need to output the items! Finally, if there are N items the memory has to be at least big enough for these and hence both memory accesses and addresses are ($O(\log N)$, pushing real behaviour back into the $O(N \log N)$ territory (although note that by similar arguments Quicksort is really $O(N (\log N)^2)$!

If you feel that these practical bit-twiddling examples feel a little contrived, there is the story of a Google employee giving a talk at Cambridge. During the presentation one of the eminent computer scientist in the audience did some quick complexity calculations in their head, and at the end stood up and said, "I like your algorithm, but unfortunately it doesn't scale". The Google employee responded, "well it works for 10 billion web pages".

4.5   What is computation?

The lower bound result for sorting is relatively rare, and, as noted, is based on information theory measures and hence works for oracles as well as 'real' computation. One of the reasons for this is that while we have had an abstract measure of information content dating back more than 70 years [21, 22], our computational metrics are, in comparison, weak.

One of my own intuitions (albeit not as informed as Hardy and Wright's!) is that some variant of Galois theory may be a way to get traction. Of course, this may simply be because the story of Évariste Galois is one of the great romances of mathematics – writing in his garret, the night before the fatal duel "there is no time, there is no time ...".

Galois theory is about what numbers it is *possible* to construct using the solution to particular equations [24] (for example square roots in geometric constructions). This is rather like non-existence proofs in computability such as the halting problem.

Of course in computing we also want to know *how many* steps it takes. While standard Galois theory does not address this, one can have variants where you are allowed only finite numbers of extension operations. The resulting sets form a tower (see Figure 8) and have some nice mathematical properties:

$$\forall a \in \mathbb{Q}_n(\sqrt{2}), b \in \mathbb{Q}_m(\sqrt{2}) \; : \; \{a+b, a-b, a*b, a/b\} \subset \mathbb{Q}_{n+m}(\sqrt{2}) \quad (9)$$

That is the sets are homomorphic to the semigroup of positive integers. If one looks at more complex Galois extensions with multiple radicals, such as $\mathbb{Q}_{n,m,s}(\sqrt{p}, \sqrt{q}, \sqrt{r})$; one ends up with a simple product semigroup if p, q and r are co-prime, but may yield more complex semigroups if they have common factors (e.g. 12, 50, 30).

As is evident this feels rather like counting computational steps of different kinds, so may be a fruitful path. I have never moved beyond this stage myself; perhaps a reader will be inspired to dig further!

5   Knowing what to model – Covid serial interval

During the summer of 2020 an estimate I made of the potential impact of university re-opening on Covid-19 deaths [10] was publicised and criticised as overstating the problem. In hindsight both later estimates by the UK Government SAGE group and actual case data in September and October showed that in fact I had been optimistic. At one point in the summer, in a BBC Radio interview, Kit Yates (University of Bath academic and popular science writer) had stated that the time between infections (called the *serial interval*) used in the paper of 3.5 days was too short and the real figure should be 5.5 days. In fact the actual modelling was independent of this figure (it just changes the time scale), but this did bring my attention to the wide variation in estimates of the serial interval.
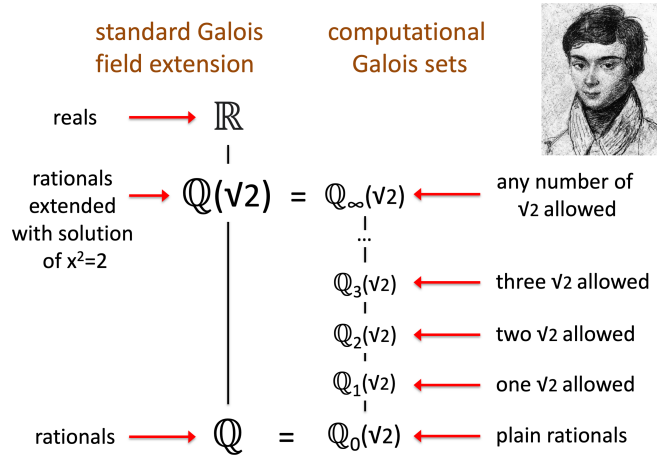
Fig. 8. Computational Galois theory

Yates was absolutely correct in that the WHO Covid-19 information at the time used a 5.5 day estimate, however, at the same time the growth graphs used by the BBC used a 3.5 day figure. Furthermore SAGE estimates of UK R factor, when compared with the doubling time, were *effectively* using a 3.5 day period (although this will have arisen out of detailed models). If one then looks further at meta-analysis papers reviewing large numbers of studies, the range of estimates varies substantially [20]. Why the discrepancy?

In part this may be due to the fact that, while R_0 and the serial interval are often stated as if they were fundamental parameters of the disease, they both depend critically on many social and environmental factors: how many contacts people have, whether indoors or outdoors, etc. In particular, R_0 tends to be higher and the serial interval shorter in densely populated areas in cold and damp climates – as is typical in the UK, but R_0 is lower and the serial interval longer in more thinly populated areas as is the case in many parts of Africa and the USA outside major cities (and even in the suburbs).

The above statement is already a qualitative–quantitative argument, but one that is perhaps so obvious it doesn't appear to be so.

A little less obvious is the complex, but comprehensible, way in which the serial interval changes when either individual caution or statutory control measures modify the spread of disease.

1. If, when infected people become symptomatic, they take substantial self-quarantining actions, this will mean less post-symptomatic contagion, but have no impact on pre-symptomatic contacts. This therefore *reduces* the serial interval.

2. For asymptomatic cases, some contacts are sporadic such as fellow passangers on public transport. For these contacts the likelihood of contagion is lower, but the average timing of those infected unlikely to be affected.

3. For asymptomatic cases, some contacts are frequent such as work colleagues and family members. For these contacts, they will have some reduction in the eventual probability of catching the disease, but crucially if they do catch it, they are likely to take longer to do so. That is, for this group the serial interval *increases*.

Note that effect (1) decreases the serial interval, effect (3) reduces it and effect (2) makes no difference. This interplay of positive and negative effects is not uncommon. One might be prompted to use further QQ-reasoning to compare the effects – it is assumed that for Covid-19 asymptomatic infections are a major driver of growth, so that might suggest (3) will be more significant than (1). Alternatively one might use the analysis to perform more detailed and precise modelling.

Finally there is a third sampling-based influence on the serial interval. Figure 9 shows the distribution of serial times for 468 infection pairs from [12]. Note the large variation: once someone is infected they may pass it on to some people straight away, but others only after a considerable period. It is the average period that is usually quoted, but this hides considerable variation.
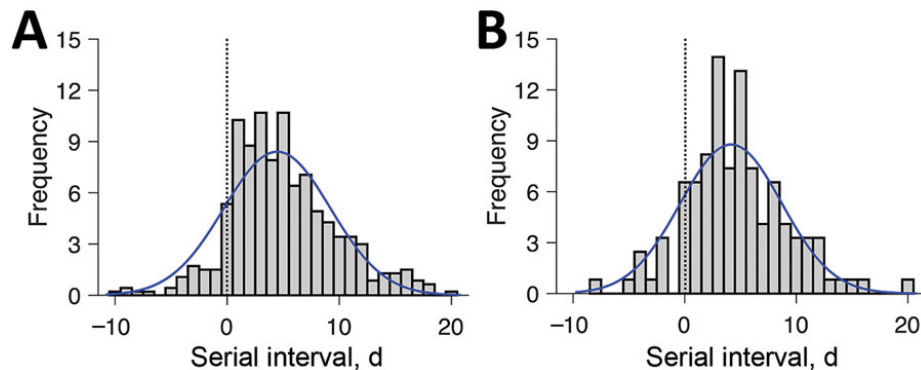
Fig. 9. Distribution of the serial interval from [12]. A is based on 468 pairs of cases and B in a subset of the 122 most reliable infection pairs. Note the negative serial intervals will be due to pre-symptomatic infections as the time measured is between the onset of symptoms of the pair.

Imagine we have perfect retrospective knowledge so that we know who caught the disease, from whom and when. There are two ways we coud measure the distribution.

1. Forward – consider at each infectious person (source), who they infect and when. This is the canonical serial interval.

2. Backward – consider at each infected person, who they were infected by (source) and how far into that source's infection.

During a period of disease growth ($R > 1$), the number of infectious people increases with time, meaning method (2) will encounter more people infected recently and hence create a shorter estimate of the serial interval than method (1). Similarly during a period of disease decline ($R < 1$) the serial interval calculated by (2) will be longer than by (1).

The serial interval combines with R to give the exponential rate of growth. If one uses the 'true' serial interval form (1) this ends up a little too large (when $R > 1$), but estimate (2) is too short. The value that gives the exponential approximation, the *effective serial interval*, is between the two.

If we wish to work out exactly how these estimates differ, we will need more precise modelling. However, the QQ-reasoning suggests what we should be modelling and directs us towards what we should be looking for in the modelling.

6   Monotonic reasoning – change at the shops and the impact of automation

Some years ago I was in a charity shop, probably buying books, I usually am. I gave the woman who was serving a ten pound note and she started to count out change – more than ten pounds of change. I told her and we worked out the right sum for the change. I think she had simply mistyped a figure into the till, but the thing that surprised me was that she hadn't noticed. This was probably due to what is often called 'automation bias', the tendency to believe what a computer tells us, even when patently wrong. Of course, automation errors, when they happen, are often gross hence the importance of being able to have a broad idea of what is a reasonable answer. In this case I was using a simple form of monotonic reasoning:

$$b > 0 \implies A - b < A \tag{10}$$

We may also do similar reasoning in two dimensions using the Poincaré property – every closed non-self intersecting line in 2D space has an inside and an outside. If you have crossed a city ring road going into the city and have not re-crossed it, then you must still be inside. However, whether this is a logical argument or more of a 'gut' knowledge about the world depends on spatial ability ... or perhaps learnt skill.

Many economic issues depend on more or less complex chains of of monotonic reasoning. Figure 10 shows two arguments for and against the value of automation. On the left hand side there is the 'pro' argument: automation leads to increased productivity, hence increases overall prosperity and this makes people better off. However, on the right-hand side is the counter argument that increased automation leads to less need for labor, hence unemployment and poverty.
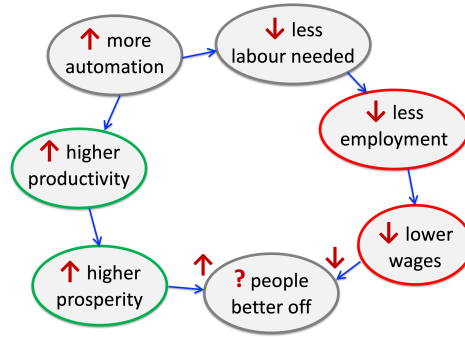
Fig. 10. Positive and negative impact of automation

Rather like the Covid examples, different arguments lead to positive and negative effects. We might resolve this by estimating (more QQ-reasoning!) the size of the effects. Perhaps more pertinently we might ask, "who benefits?"

Laying out an argument in this way also makes it easier to debate the steps in the inferences, rather like argumentation systems such as IBIS (issue-based information system) [16, 3]. For example, the link that suggests that automation leads to less labour has been questioned using the example of Amazon, which in 2016 installed 15,000 robots, but instead of reducing labour in fact also increased their employees by 46% [13]. This has then been used to argue that robots increase employment [14]. However, it is likely that the growth is due to the left hand thread in Figure 11: robots, improved competitiveness, helped the company grow its sales and hence increased employment *at Amazon*. Seeing this immediately brings to mind the right hand arc of the same figure, that the growth of Amazon has probably shrunk other businesses and hence decreased employment elsewhere.
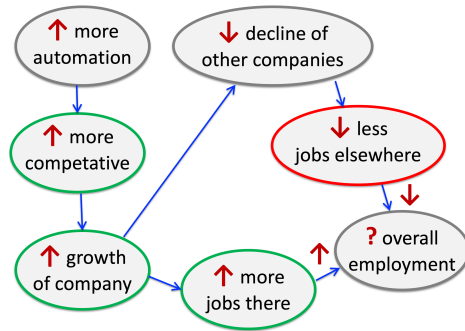

Fig. 11. Does automation create jobs?

7   Formalising and visualising QQ – Allen's Interval Calculus

Figures 10 and 11 are both a visualisation of the argument and also a type of formal representation of the qualitative–quantitative reasoning about automation. It a form of high-level argumentation, similar to safety cases used in the nuclear and aviation industry. While the validity of each judgement step ('this increases that') is a human one, given such lower-level judgements, the overall reasoning can be verified:

> **given**
>   *increase in A leads to an increase in B*   (human judgement)
> **and**
>   *increase in B leads to an increase in C*   (human judgement)
> **conclude**
>   *increase in A leads to an increase in C*   (formal inference)

We can find other examples of formalisation of QQ in the literature. Some force you to make the informal judgements very precise, for example fuzzy logic demands a precise shape for the uncertainty function and Bayesian statistics require that you encode your belief as if it were a probability [11]. Other methods embrace the human-like reasoning more wholeheartedly, including various representations of naïve physics or informal reasoning used in cognitive science and artificial intelligence such as Allen's Interval Calculus [1] for reasoning about temporal events (see Figure 12).
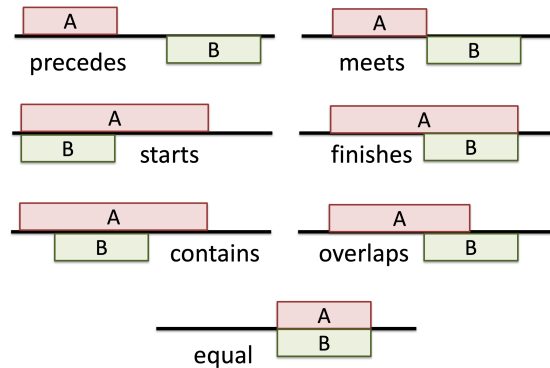
Fig. 12. Different temporal relations (from [1])

8   Discussion and call to action

We have seen a variety of examples of qualitative–quantitative reasoning. Some were about gaining informal understanding from formal or theoretical models; some were about rough sizes: montonicity and orders of magnitude; and some were about numerical modelling: how to guide what we model and how to turn idealised or simplified models into representations that are useful for decision making. While many of the examples were about academic or professional use, others were about the general populous. Indeed, in a data-dominated world, understanding numerical arguments is essential for effective citizenship.

We have also seen that there are existing methods and representations to help with qualitative–quantitative reasoning, but relatively few given the criticality in so many walks of life.

There are three lessons I'd like the reader to take away:

- recognise when you are you using qualitative–quantitative reasoning so that you can think more clearly about your own work, and perhaps make it more accessible or practically useful
- realise that it is a potential area to study theoretically in itself – are there ways to formalise or visualise some of the informal reasoning we use about formal things!
- seek methods and tools to help others think more clearly about this: in universities, industry and schools.

References

1. Allen, J.F.: Maintaining knowledge about temporal intervals. Commun. ACM **26**(11), 832–843 (Nov 1983). https://doi.org/10.1145/182.358434, https://doi.org/10.1145/182.358434
2. Campos, J.C., Fayollas, C., Harrison, M.D., Martinie, C., Masci, P., Palanque, P.: Supporting the analysis of safety critical user interfaces: An exploration of three formal tools. ACM Trans. Comput.-Hum. Interact. **27**(5), 1–48 (Aug 2020). https://doi.org/10.1145/3404199, https://doi.org/10.1145/3404199
3. Conklin, J., Begeman, M.L.: Gibis: A hypertext tool for exploratory policy discussion. ACM Trans. Inf. Syst. **6**(4), 303–331 (Oct 1988). https://doi.org/10.1145/58566.59297, https://doi.org/10.1145/58566.59297
4. Conway, J.H.: On numbers and games. Academic Press (1976)
5. Dantzig, G.B.: Origins of the Simplex Method, p. 141–151. Association for Computing Machinery, New York, NY, USA (1990), https://doi.org/10.1145/87252.88081
6. Dix, A., Mancini, R., Levialdi, S.: The cube – extending systems for undo. In: Proceedings of DSVIS'97. pp. 473–495. Eurographics (1997), https://alandix.com/academic/papers/dsvis97/
7. Dix, A.J., Marchant, J.A.: A mathematical model of the transport and deposition of charged spray drops. Journal of Agricultural Engineering Research **30**, 91–100 (1984)
8. Dix, A.J., Runciman, C.: Abstract models of interactive systems. In: Johnson, P., Cook, S. (eds.) People and Computers: Designing the Interface. pp. 13–22. Cambridge University Press (1985), https://alandix.com/academic/papers/PIE85/PIE-paper.html

9. Dix, A.: Formal Methods for Interactive Systems. Academic Press (1991), https://alandix.com/books/formal/

10. Dix, A.: Impact of a small number of large bubbles on covid-19 transmission within universities (2020), https://arxiv.org/abs/2008.08147, arXiv 2008.08147

11. Dix, A.: Bayesian Statistics, chap. 2, p. 141–151. Cambridge University Press (2021)

12. Du, Z., Xu, X., Wu, Y., Wang, L., Cowling, B.J., Meyers, L.A.: The serial interval of covid-19 from publicly reported confirmed cases. medRxiv (2020). https://doi.org/10.1101/2020.02.19.20025452, https://www.medrxiv.org/content/early/2020/03/13/2020.02.19.20025452

13. González, A.: Amazon's robot army grows by 50 percent. The Seatle Times (Dec 2016), https://www.seattletimes.com/business/amazon/amazons-robot-army-grows/

14. Hamid, O.H., Smith, N.L., Barzanji, A.: Automation, per se, is not job elimination: How artificial intelligence forwards cooperative human-machine coexistence. In: 2017 IEEE 15th International Conference on Industrial Informatics (INDIN). pp. 899–904 (2017). https://doi.org/10.1109/INDIN.2017.8104891

15. Hardy, G.H., Wright, E.M.: An Introduction to the Theory of Numbers. Oxford, fourth edn. (1975)

16. Kunz, W., Rittel, H.W.J., Messrs, W., Dehlinger, H., Mann, T., Protzen, J.J.: Issues as elements of information systems. Tech. rep., University of California, Berkeley (1970), https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.134.1741

17. Locke, J.: Second Treatise of Government (1689), https://gutenberg.org/ebooks/7370

18. Mancini, R.: Modelling Interactive Computing by Exploiting the Undo. Dottorato di ricerca in informatica, ix-97-5, Università degli Studi di Roma "La Sapienza" (1997), https://hcibook.net/people/Roberta/

19. Pearce, F.: Earth's most important rivers are in the sky – and they're drying up. New Scientist **3254** (Nov 2019), https://www.newscientist.com/article/mg24432540-600-earths-most-important-rivers-are-in-the-sky-and-theyre-drying-up/

20. Rai, B., Shukla, A., Dwivedi, L.: Estimates of serial interval for covid-19: A systematic review and meta-analysis. Clin Epidemiol Glob Health pp. 157–161 (2021). https://doi.org/10.1016/j.cegh.2020.08.007

21. Shannon, C.E.: A mathematical theory of communication. The Bell System Technical Journal **27**(3), 379–423 (1948). https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

22. Shannon, C.E., Weaver, W.: The Mathematical Theory of Communication. University of Illinois Press, Urbana, IL (1949)

23. Shirriff, K.: Ibm punched card sorter type 83, from 1955, https://www.youtube.com/watch?v=AyO3n20SpoI

24. Stewart, I.N.: Galois Theory (4th ed.). Chapman and Hall/CRC (2015), https://doi.org/10.1201/b18187, (First ed. 1973)

25. Thimbleby, H.: Generative user-engineering principles for user interface design. In: INTERACT '84. pp. 661–666 (1984)