Modelling Complex Adaptive Systems: Toy or Tool?

— Robert Marks, MBS and UNSW.

Agent-based models can be characterised as Complex Adaptive Systems, and can model aspects of real-world CAS.

ABM are not the only CAS models — non-linear models such as David has shown you can also generate some CAS behaviour such as:

- **deterministic chaos (extreme sensitivity to initial conditions),**
- **strange attractors, etc**

But we need *multiple agents* to derive such CAS characteristics as:

- **emergence at a higher (macro) level,**
- **self-organisation.**

Toys?

We all remember the "Sim" line of computer games — an early example of ABM.

Toys like this are serious business:

although not an ABM, last year *Call of Duty: Black Ops* took in over USD$650 million of sales in the game's first five days, which set a five-day global record for a movie, book or videogame.

Tools?

Despite the general limitation of simulations in general (and ABM in particular) to provide sufficiency only,

Simulations can help explain simple phenomena in principle, and complex historical phenomena — although this latter requires validation.

One use of CAS modelling is to find *sufficient conditions* for the model to exhibit such characteristics — I gather that's what you're doing with your team projects.

And you may derive suggestive trade-offs and hypotheses, to be tested.

Two Kinds of ABM

I suggest that we can think of two kinds of ABM:

1. ***demonstrative* ABM models**

   These models demonstrate principles, rather than tracking historical phenomena. A demonstrative ABM is an existence proof.
   Examples: Schelling's Segregation Game, Boys and Girls, the Emergence of Risk Neutrality:
   `http://www.agsm.edu.au/bobm/papers/ralet.pdf`

2. ***descriptive* ABM models.**

   These models attempt to derive sufficient conditions to match historical phenomena, as reflected in historical data. This requires validation.
   Examples: Midgley et al. modelling brand rivalry, the poli sci models (David).

Proofs of Sufficiency

Simulation models are proofs of sufficiency — "with this model and these parameters it is possible to obtain this behaviour."

Closed-form models in general can also prove necessity — "with this model and these parameters or with that model and those parameters it is possible to obtain this behaviour and with no other models or parameters."

ABM stockmarkets attempt to prove the existence of models (as simple as possible but no simpler) that exhibit such real-world phenomena as fat-tailed distributions, volatility clustering, etc.

It was the ability of LeBaron et al.'s artificial stock market that must have led Jean-Claud Trichet to suggest agent-based models as a way of augmenting — not surplanting — existing macroeconomic models.

Simulation and Necessity?

Mathematical "model A" comprises the conjunction $(a_1 \wedge a_2 \wedge a_3 \cdots \wedge a_n)$, where $\wedge$ means "AND", and the a_i denote the elements (equations, parameters, initial conditions, etc) that constitute the model.

***Sufficiency*: If model A exhibits the desired target behaviour B, then model A is sufficient to obtain exhibited behaviour B: $A \Rightarrow B$**

Thus, any model that exhibits the desired behaviour is sufficient, and demonstrates one conjunction of elements (or one model) under which the behaviour can be simulated.

But if there are several such models, how can we choose among them? And what is the set of all such conjunctions (models)?

Necessity

Necessity: **Only those models A belonging to the set of necessary models $\mathcal{N}$ exhibit target behaviour B.**

That is, $(A \in \mathcal{N}) \Rightarrow B$, and $(D \notin \mathcal{N}) \not\Rightarrow B$.

A difficult challenge: determine the set of necessary models, $\mathcal{N}$.

Since each model $A = (a_1 \wedge a_2 \wedge a_3 \cdots \wedge a_n)$, searching for the set $\mathcal{N}$ of necessary models means searching in a high-dimensional space, with no guarantee of continuity, and a possible large number of non-linear interactions among elements.

Lack of Necessity Means ...

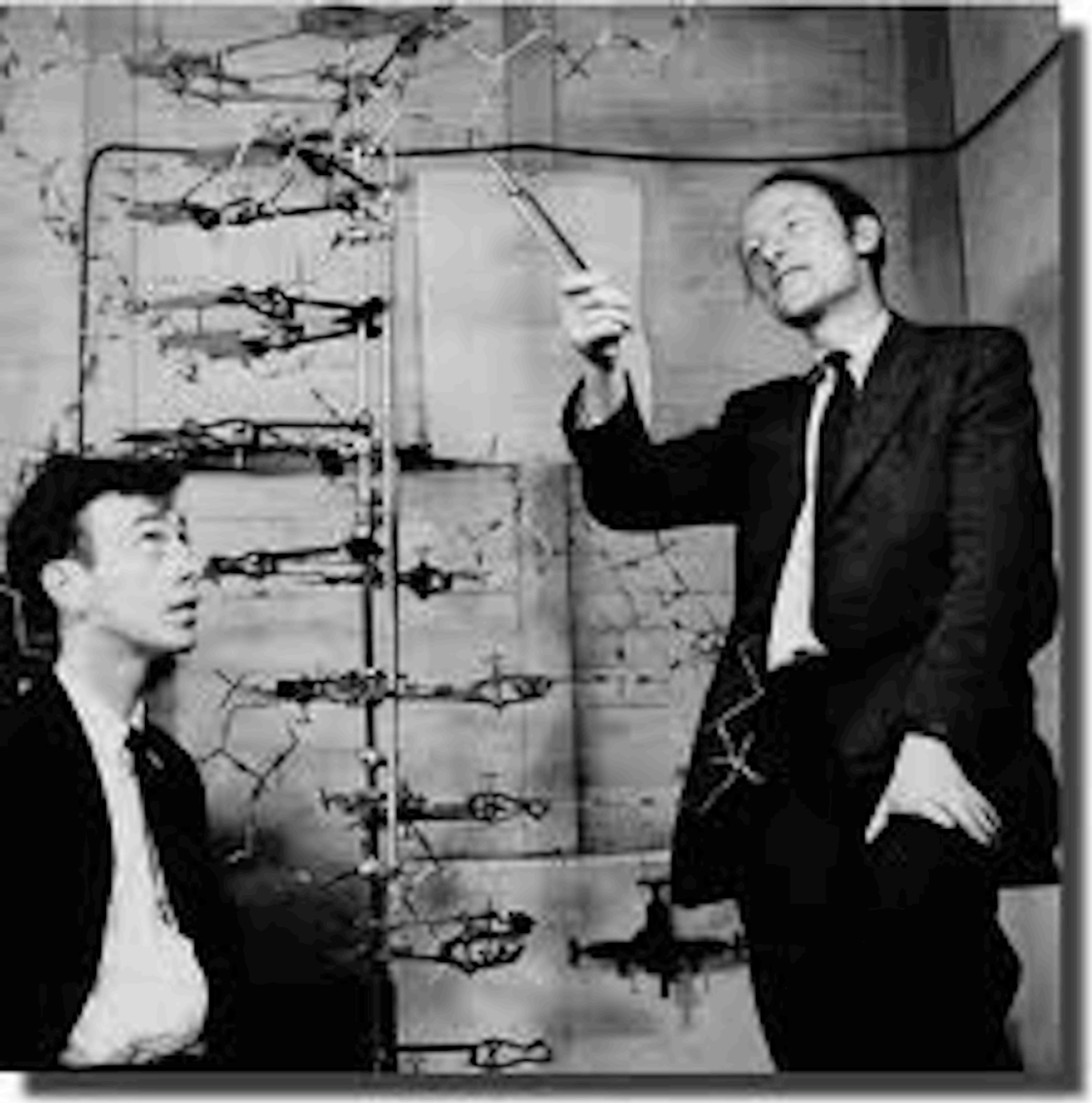
For instance, if $D \not\Rightarrow B$, it does not mean that *all* elements a_i of model D are invalid or wrong, only their conjunction, that is, model D.

It might be only a single element that precludes model D exhibiting behaviour B.

But determining whether this is so and which is the offending element is a costly exercise, in general, for the simulator.

Without clear knowledge of the boundaries of the set $\mathcal{N}$ of necessary models, it is difficult to generalise from simulations.

Only when the set $\mathcal{N}$ of necessary models is known to be small (such as in the case of DNA structure by 1953 when Watson & Crick were searching for it) is it relatively easy to use simulation to derive necessity.



MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

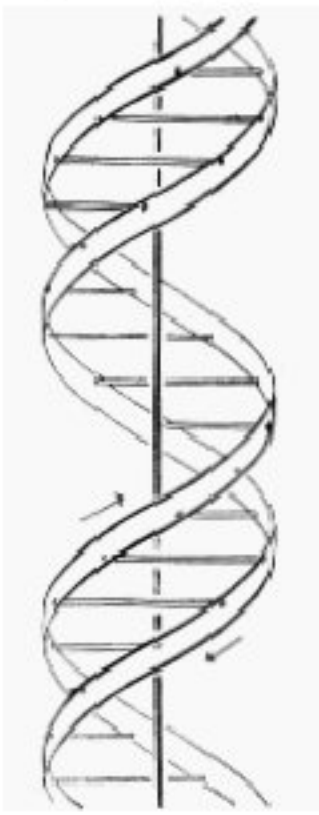
WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three inter-twined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion. this structure is unsatisfactory for two reasons:

(1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions. namely, that each chain consists of phosphate diester groups joining ß-D-deoxy-ribofuranose residues with 3',5' linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow righthanded helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions.



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the cbains together. The vertical line marks the fibre axis

Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3-4 A. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain, does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain, is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in time following, communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereo-chemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. *M.* H. F. Wilkins, Dr. R. E. Franklin and their co-workers at King's College, London. One of us (J.D.W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J.D. WATSON
F.H. C. CRICK

Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge. April 2.

[1] Pauling, L., and Corey, R. B. nature, 171, 346 (1953); Proc, U.S. Nat Acad. Sci., 39, 84 (1953).

[2] Furberg, S., Acta Chem. Scand., 6, 634 (1952).

[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., Biochim, et Biophys. Acta, 9402 (1952).

[4] Wyatt, G.R. J. Gen. Physiol., 36 201 (1952).

[5] Astbury, W.T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947)

[6] Wilkins, M. H. F. and Randall, J. T. Biochim. et. Biophys. Acta, 10, 102 (1953).

Families of Simulation Models

1. **System Dynamics SD**
   (from differential equations)

2. **Cellular Automata CA**
   (from von Neumann & Ulam, related to Game Theory)

3. **Multi-Agent Models MAM, or Agent-Based Computational Economics ACE, or Agent-Based Models ABM, or Multi-Agent Systems MAS**
   (from Artificial Intelligence)

4. **Learning Models LM**
   (from Simulated Evolution and from Psychology)

Comparison of Simulation Techniques

Gilbert & Troitzsch (2005) compare these (and others):

Technique	Number of Levels	Communication between agents	Complexity of agents	Number of agents
SD	1	No	Low	1
CA	2+	Maybe	Low	Many
MAM	2+	Yes	High	Few
LM	2+	Maybe	High	Many

Number of Levels: "2+" means the technique can model more than a single level (the individual, or the society) and the interaction between levels.

This is necessary for investigating emergent phenomena.

So "agent-based models" excludes simple Systems Dynamics (SD) models, but can include the others.

As Simple as Possible … but no Simpler.

As ABM modellers: remember to simplify — parsimony is the watchword.

Why?
Great freedom to make whatever assumptions you like with ABM (unlike the assumptions embodied in closed-form math models), so …

Don't keep adding bells and whistles just because you can — it will confuse your audience, will give leverage to the skeptics, and will exponentially burden your sensitivity anaysis.

Add diagnostics but don't complicate your model for its own sake if your model is demonstrative, or even if it is descriptive.

You can still derive trade-offs and statistics, as David reported with the poli sci models.