Model theory and modeling

Robert J. Carroll

Last revised: 5 May 2026

---

1 Motivation

Political scientists and economists use the word *model* a lot. "A model of voting behavior." "The Cournot model of duopoly." "The agenda-setter model." What is one actually doing when one writes down a model of this kind? The standard answer in working practice goes something like: a model is a setup of agents (with action sets, preferences, information), an environment (states, signals, payoffs), and a notion of how outcomes are determined. One then proves things about the model — existence of equilibrium, the form of the equilibrium strategies, comparative statics in parameters, and so on.

This is structurally close to the formal-logic vocabulary of the previous handout: a *structure* is a domain plus interpretations of the symbols of a signature, and "writing down a model" is, near enough, exhibiting a structure. The connection is more than an analogy. The philosophical thesis known as the *semantic view* of theories (Patrick Suppes, Bas van Fraassen, and others) is that scientific theories are best understood as classes of models in essentially this sense, rather than as syntactic axiomatizations. For political science the semantic view is not just plausible; it is what working modelers already do, with the formal-logic vocabulary providing the precise version of the everyday usage.

This handout develops the basic model-theoretic apparatus — theories as sets of sentences, the class of models of a theory, elementary equivalence, isomorphism — and then steps back to the semantic view of theories and how it lines up with applied modeling. It rounds out the foundations sequence: from here the project moves on to applications (decision theory, then onward), with the formal-logic foundations in place and an account of what "a model" is in this enterprise.

2 Theories and their models

We have signatures and structures from the previous handout. The third basic object is a *theory*.

Definition 1. A *theory* in a signature σ is a set T of σ-sentences. The *models* of T, written $\mathrm{Mod}(T)$, is the class of σ-structures that satisfy every sentence in T:

$$\mathrm{Mod}(T) = \{\mathfrak{A} : \mathfrak{A} \models \varphi \text{ for every } \varphi \in T\}.$$

We say T is *satisfiable* if $\mathrm{Mod}(T) \neq \emptyset$, and *consistent* (in the syntactic sense) if $T \nvdash \bot$.

By soundness and completeness of FOL (Gödel 1930), satisfiability and consistency coincide — the syntactic notion of consistency exactly captures the semantic existence of a model. We will not prove this; the result of the previous handout shoulders the work.

Example 2 (Theory of a preference relation)**.** Let σ_{pref} have a binary relation symbol $\succsim$. Let T_{cpo} consist of:

$$\begin{aligned} &\forall x\,(x \succsim x) && \text{(reflexivity)}\\ &\forall x\forall y\forall z\,((x \succsim y \wedge y \succsim z) \to x \succsim z) && \text{(transitivity)}\\ &\forall x\forall y\,(x \succsim y \vee y \succsim x) && \text{(completeness)} \end{aligned}$$

Then $\mathrm{Mod}(T_{\text{cpo}})$ is the class of all complete preorders — exactly the relations that the order-theory handout took as the standard preference structure. Every individual "preference profile" in a model is a member of this class; theorems about complete preorders (e.g., representability by a utility function under further hypotheses) are statements about all of $\mathrm{Mod}(T_{\text{cpo}})$.

Example 3 (Theory of a single-peaked profile, sketch)**.** Take a many-sorted signature with sorts V (voters), A (alternatives), and $\mathbb{R}$ (real numbers, to host ideal points), with: a function $i : V \to \mathbb{R}$ assigning each voter an ideal point; a function $p : A \to \mathbb{R}$ assigning each alternative its policy position; and the relation $\succsim$ of sort (V, A, A). Take T_{sp} to consist of (informally stated): a voter's preference between two alternatives is determined by which is closer to their ideal point. The class $\mathrm{Mod}(T_{\text{sp}})$ is exactly the family of single-peaked-on-the-line preference profiles — the input domain of the median voter theorem. The MVT is then a sentence true in every member of $\mathrm{Mod}(T_{\text{sp}})$ subject to the further axiom that there is an odd number of voters, etc.

The map $T \mapsto \mathrm{Mod}(T)$ has a natural inverse direction: starting from a class of structures, what sentences are true in all of them?

Definition 4. For a class $\mathcal{K}$ of σ-structures, the *theory of* $\mathcal{K}$ is

$$\mathrm{Th}(\mathcal{K}) = \{\varphi : \mathfrak{A} \models \varphi \text{ for every } \mathfrak{A} \in \mathcal{K}\}.$$

The two operations Mod and Th are order-reversing in the natural way: if $T \subseteq T'$, then $\mathrm{Mod}(T) \supseteq \mathrm{Mod}(T')$ (more axioms means fewer models); if $\mathcal{K} \subseteq \mathcal{K}'$, then $\mathrm{Th}(\mathcal{K}) \supseteq \mathrm{Th}(\mathcal{K}')$ (fewer structures means more sentences true throughout). Composing them gives *deductive closure*: $\mathrm{Th}(\mathrm{Mod}(T))$ is the set of all sentences T semantically entails, which by completeness equals the syntactic deductive closure of T.

3 Elementary equivalence and isomorphism

When are two structures "the same"? The strongest answer is *isomorphism*: a bijection between the universes that respects the signature exactly. The weaker answer is *elementary equivalence*: indistinguishable by any first-order property.

Definition 5. Two σ-structures $\mathfrak{A}$ and $\mathfrak{B}$ are:

- *isomorphic* ($\mathfrak{A} \cong \mathfrak{B}$) if there is a bijection $h : A \to B$ such that for every constant c, $h(c^{\mathfrak{A}}) = c^{\mathfrak{B}}$; for every function symbol f, $h(f^{\mathfrak{A}}(\vec{a})) = f^{\mathfrak{B}}(h(\vec{a}))$; and for every relation symbol R, $\vec{a} \in R^{\mathfrak{A}}$ iff $h(\vec{a}) \in R^{\mathfrak{B}}$;
- *elementarily equivalent* ($\mathfrak{A} \equiv \mathfrak{B}$) if for every σ-sentence φ, $\mathfrak{A} \models \varphi$ iff $\mathfrak{B} \models \varphi$.

Proposition 6. $\mathfrak{A} \cong \mathfrak{B}$ *implies* $\mathfrak{A} \equiv \mathfrak{B}$.

Proof sketch. By induction on the structure of φ. An isomorphism h preserves the value of every term and the membership of every tuple in every relation, hence preserves the satisfaction of every atomic formula. Boolean connectives and quantifiers preserve under isomorphism by direct verification. The full claim is then that $\mathfrak{A} \models \varphi[s]$ iff $\mathfrak{B} \models \varphi[h \circ s]$, with the sentence case being where the assignment drops out. □

The converse is generally false: two structures can be elementarily equivalent without being isomorphic. The standard examples are infinite. Two countably infinite dense linear orders without endpoints are isomorphic (Cantor's theorem; recall the cardinality handout's exercise on back-and-forth). But a countable and an uncountable dense linear order without endpoints are elementarily equivalent yet not isomorphic — $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ are an instance.[1]

For applied modeling, the practical concept is usually elementary equivalence rather than isomorphism. Two political-economy models are "the same" in any usable sense if they make all the same first-order statements true — if no first-order property distinguishes them. Isomorphism is a stronger notion that pins down the structures element-by-element, and is rarely the relevant question.

A theory T is *complete* if for every sentence φ in its signature, either $T \models \varphi$ or $T \models \neg\varphi$.[2] Complete theories have the property that all their models are elementarily equivalent (a sentence is in T iff its negation is not, and structures satisfying T all decide every sentence the same way). Conversely, a consistent theory whose models are pairwise elementarily equivalent is complete. "Decidable up to elementary equivalence" is what completeness of a theory amounts to.

4 The semantic view of theories

So far we have an apparatus: theories as sets of sentences, structures as their models, elementary equivalence as the natural notion of indiscernibility. The philosophical thesis we now want to flag is that this apparatus is the right way to think about theories *in general*, including informally presented theories in the working empirical sciences — and that for political science and economics in particular, the apparatus matches everyday practice essentially exactly.

The contrast is between two pictures of what a scientific theory is.

[1] The phenomenon is essentially a cardinality phenomenon, and it is what the *Löwenheim–Skolem theorem* formalizes: every consistent first-order theory in a countable language has models of every infinite cardinality. The downward direction (every infinite model has a countable elementary substructure) was Löwenheim's original 1915 result; the upward direction (models of arbitrarily large cardinality exist) was Skolem and Tarski's. The most striking consequence is *Skolem's paradox*: ZF set theory, formulated in a countable first-order language, has a countable model. The set theorist's "the set of all real numbers" is uncountable, but the model in which the theorem is interpreted is itself countable, with "uncountable" meaning "no bijection-with-$\mathbb{N}$ *exists in this particular model.*" This is paradoxical only on a naive reading; once one separates "cardinality as a structural property of a model" from "cardinality as a property of an external set," the paradox dissolves into a deep observation about the relationship between syntactic theories and their semantic interpretations. For political-economy work the moral is small but real: a first-order theory does not pin down a unique structure, and "the model" a theorist has in mind is always a particular structure, not the whole class of models satisfying the same axioms. The class is the theorem-prover's playground; the structure is the modeler's. Marker (2002) treats Löwenheim–Skolem and the basic model-theoretic apparatus carefully.

[2] The word *complete* now appears in three distinct technical senses across this project: (1) the LUB property of $\mathbb{R}$ (analysis), (2) Gödel's completeness theorem $\models \Leftrightarrow \vdash$ for FOL (proof theory), and (3) "a theory pins down all sentences as true or false" (model theory). They are all important and they are all standard. Context disambiguates: a complete preorder is a relation, a complete proof system is a system, a complete theory is a theory. The overload is unfortunate but unavoidable.

The syntactic view (sometimes called the *received view*, associated with Carnap and the logical empiricists in the 1930s–50s) treats a theory as a formal language together with a list of axioms, with consequences obtained by syntactic deduction. A theory *is* its axiomatization; "Newtonian mechanics" is the deductive closure of Newton's three laws (suitably formalized) plus auxiliary mathematical apparatus.

The semantic view (Suppes 1957, 1967, 2002; van Fraassen 1980; Sneed 1971; Stegmüller's structuralist programme) treats a theory as a class of *models*, where "model" here means a set-theoretic structure of the kind we have developed in the previous handout. The axioms, when they exist, are devices for picking out the class — but the theory *is* the class, not the axiomatization. "Newtonian mechanics" on this view is the class of structures (consisting of point particles, position-and-velocity functions, force functions, etc.) that satisfy Newton's laws.

The two views agree in formal content for first-order theories: by completeness, $T \models \varphi$ iff $T \vdash \varphi$, so the deductive closure of the axioms is exactly the theory of the class of models. The disagreement is about which side of the equivalence is conceptually primary, and the practical consequences differ once one gets outside the strictly first-order case (where higher-order or set-theoretic apparatus is needed) or once one cares about the relationship between formal theories and the empirical situations they purport to model.

For political science and economics the semantic view is a much closer fit to working practice. When a theorist writes down "a Cournot duopoly," they specify: a set of firms (here, two), an inverse demand function $p(Q) = a - bQ$, a cost function for each firm, payoff functions, a strategy space (quantities), and the equilibrium concept (Nash). This is not a syntactic axiomatization. It is a many-sorted structure — carriers (firms, real numbers), distinguished functions (p, u_i), distinguished elements (parameter values a, b, c). When a different theorist writes "a Bertrand duopoly," a different structure is exhibited (price competition rather than quantity competition). When a textbook says "the Cournot model," it is referring to a class of such structures, parameterized by the inverse demand and cost functions allowed.

What unifies the political-economy modeling tradition with the formal-logic apparatus is that "model = structure." The everyday use of "model" is essentially the formal-logic "structure"; the everyday use of "the family of models of this kind" is essentially the formal-logic "$\mathrm{Mod}(T)$." The semantic view of theories, then, is more than a philosophical thesis — for political-economy modeling it is a description of what is already happening, with a precise vocabulary attached.

A few cautionary points. First, the semantic view does not say syntactic apparatus is unimportant; it says the syntactic apparatus is in service of the structures, not the other way around. When one wants to prove an impossibility theorem (Arrow, Gibbard–Satterthwaite, Sen), one does so by syntactic argument over the axioms — the syntactic side is doing the heavy lifting. The semantic view says only that the *point* of the syntactic argument is to constrain the class of structures.

Second, the model-theoretic "model" should not be confused with the philosophy-of-science usage of "model" as in "a scale model," "a toy model," "a model of a hurricane." The latter usage is about *representation* of empirical phenomena, while the formal-logic usage is about *satisfaction* of axioms. For political-economy modeling the two usages bleed into each other: a Cournot model is both a satisfaction of certain axioms and a representation of (a stylization of) duopoly markets. The slippage is mostly harmless but worth flagging when someone seems to mean one thing by "model" while their interlocutor means the other.

Third, the semantic view does not eliminate the work of writing down axioms or making structures rigorous — it relocates it. The more concrete one's model is (a specific Cournot setup, with parameters specified), the more it is a single structure; the more general the modeling claim ("the Cournot framework predicts X under conditions Y"), the more it is a quantification over a class of structures. The shift to the semantic view sharpens which kind of claim one is making in a given moment of work.

5 What's next

This handout closes the foundations sequence. The reader equipped with logic (propositional and first-order), proof systems and the soundness–completeness bridge, basic set theory, order theory, the analysis cluster, and now model theory has the formal vocabulary to read essentially any paper in formal political-economy theory and to follow the rigorous content of any reasonable applied-modeling paper. The applications clusters — decision theory, then onward — pick up from here.

For deeper development of the model-theoretic apparatus we have only sketched, see Marker (2002) (a clean modern textbook focused on stability and the more analytic strands of the subject) or Hodges (1997) (the standard reference on classical model theory). For the philosophical literature on the semantic view, Suppes (2002) is the late-career Suppes summa, Fraassen (1980) is the canonical statement of the constructive empiricist version, and Frigg and Hartmann (2020) is a broad survey of the philosophy-of-science literature on models in general. The intersection between formal model theory and applied economic modeling — between Marker's textbook and the Cournot-model worldview — has been developed less than one might expect; this handout's case for the semantic view is an attempt to flag the affinity.

6 Exercises

Exercise 7. Identify $\mathrm{Mod}(T)$ for each of the following theories. (You may state the answer informally.)

1. $T_1 = \{\forall x\,(x \succsim x),\ \forall x \forall y\,(x \succsim y \to y \succsim x),\ \forall x \forall y \forall z\,((x \succsim y \wedge y \succsim z) \to x \succsim z)\}$.
2. $T_2 = T_1 \cup \{\forall x \forall y\,(x \succsim y \to x = y)\}$.
3. $T_3 = \{\forall x \forall y\,(x \succ y \to \neg y \succ x),\ \forall x \forall y \forall z\,((x \succ y \wedge y \succ z) \to x \succ z)\}$.

Exercise 8. Verify the order-reversing properties: if $T \subseteq T'$, then $\mathrm{Mod}(T) \supseteq \mathrm{Mod}(T')$; if $\mathcal{K} \subseteq \mathcal{K}'$, then $\mathrm{Th}(\mathcal{K}) \supseteq \mathrm{Th}(\mathcal{K}')$. Conclude that $T \subseteq \mathrm{Th}(\mathrm{Mod}(T))$ and $\mathcal{K} \subseteq \mathrm{Mod}(\mathrm{Th}(\mathcal{K}))$. (When does equality hold in each?)

Exercise 9. Prove Proposition 6 carefully: an isomorphism preserves satisfaction of every formula. (Induct on the structure of φ, with the assignment s as a free parameter.)

Exercise 10. Show that $(\mathbb{Z}, <)$ and $(\mathbb{Q}, <)$ as ordered structures are *not* elementarily equivalent. (Hint: write a sentence that is true in one but not the other. The integers have a successor function; the rationals do not.)

Exercise 11. A theory T is *complete* if for every sentence φ, either $T \models \varphi$ or $T \models \neg\varphi$. Show: if T is complete and consistent, then any two models $\mathfrak{A}, \mathfrak{B} \models T$ are elementarily equivalent. Conversely, if all models of a consistent theory are pairwise elementarily equivalent, then the theory is complete.

Exercise 12 (A Cournot signature)**.** Specify a many-sorted signature appropriate for a two-firm Cournot duopoly. (Suggested sorts: firms, quantities (real numbers), prices (real numbers). Suggested function symbols: an inverse demand function p, a cost function c_i for each firm, a payoff function π_i. Suggested distinguished elements: the parameters of the demand function.) Then write the sentence "firm i chooses the quantity that maximizes its payoff given firm j's choice" — the best-response condition — as a many-sorted first-order sentence over your signature.

Exercise 13 (Median voter theorem as a model-theoretic claim)**.** Sketch a many-sorted signature for the standard Black/Downs setup: voters, alternatives, ideal points, single-peaked preferences. State the median voter theorem ("the median voter's ideal point is a Condorcet winner under majority rule, when the number of voters is odd and preferences are single-peaked") as a model-theoretic claim of the form "for every structure satisfying [axioms], [conclusion sentence] holds." You do not need to prove the theorem; the exercise is in articulating its model-theoretic shape.

Exercise 14. Take the theory T_{cpo} of complete preorders from Example 2. Exhibit two finite models of T_{cpo} that are not isomorphic. Are they elementarily equivalent? (For finite structures, elementary equivalence and isomorphism coincide — a fact you may use; spell out why it makes the answer to the second question clean.)

Exercise 15 (Arrow's theorem as a non-existence claim)**.** Arrow's impossibility theorem is conventionally stated as: there is no social welfare function (mapping profiles of individual preference rankings to a social preference ranking) satisfying universal domain, transitivity, the Pareto condition, independence of irrelevant alternatives, and non-dictatorship, when there are at least three alternatives and at least two voters. Sketch the corresponding many-sorted signature and the axioms; the theorem then takes the form "$\text{Mod}(T_{\text{Arrow}}) = \emptyset$" (the theory has no models). What is the political-science content of articulating it this way? What does it not change about the theorem itself?

Exercise 16 (Conceptual)**.** The semantic view of theories says that "the Cournot model" is a class of structures rather than a syntactic axiomatization. Identify a specific paper (or textbook treatment) of any political-economy model with which you are familiar, and decide whether the paper is best understood semantically (as exhibiting a class of structures) or syntactically (as an axiomatic deduction). Two paragraphs. The point of the exercise is to test the thesis on a concrete case rather than in the abstract.

References

Fraassen, Bas C. van (1980). *The Scientific Image*. Oxford: Oxford University Press.

Frigg, Roman and Stephan Hartmann (2020). *Models in Science*. The Stanford Encyclopedia of Philosophy, Spring 2020 Edition, edited by Edward N. Zalta. URL: https://plato.stanford.edu/archives/spr2020/entries/models-science/.

Hodges, Wilfrid (1997). *A Shorter Model Theory*. Cambridge: Cambridge University Press.

Marker, David (2002). *Model Theory: An Introduction*. New York: Springer.

Suppes, Patrick (2002). *Representation and Invariance of Scientific Structures*. Stanford: CSLI Publications.