Berliant, Marcus; Fujita, Masahisa

Conference Paper

The fine microstructure of knowledge creation dynamics: reporting further advances

54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia

Provided in Cooperation with:
European Regional Science Association (ERSA)

The Fine Microstructure of Knowledge Creation Dynamics*

(Work in Progress)

Marcus Berliant** and Masahisa Fujita±

July 20, 2014

Abstract

This paper presents a new framework for modeling the fine microstructure of knowledge creation dynamics. Our focus is on the creation of working knowledge used in innovation, for example, the knowledge used by a researcher in the economics profession. The framework has been developed to address the following questions: What is the appropriate way to model the operational structure of working knowledge? How are specific new ideas, research papers, and patents created by a research worker or a group of them, based on the current stock of knowledge? What roles do dynamics, heterogeneity of ideas, heterogeneity of researchers, and cities or regions play? Using our framework, first we study how a researcher creates a new literature, choosing new assumptions, models, implications, and observations in each step. Then we suggest how to extend the analysis to the N-person case in multiple cities or regions.

JEL Classification Numbers: Keywords: Knowledge creation; Dynamic process of innovation; Reverse engineering; Chu space; Formal concept analysis

*The authors thank Isabella Prindable for helpful comments. The first author is grateful for funding from Konan University. The second author is grateful for Grants Aid for Scientific Research Grant A 26245037 from the Japanese Ministry of Education and Science. Evidently, the authors alone are responsible for any remaining errors and for the views expressed herein.

**Department of Economics, Washington University, Campus Box 1208, 1 Brookings Drive, St. Louis, MO 63130-4899 USA. Phone: (1-314) 935-8486, Fax: (1-314) 935-4156, e-mail: berliant@artsci.wustl.edu

±RIETI, Research Institute of Economy, Trade and Industry, 1-3-1 Kasumigaseki, Chiyoda-ku, Tokyo, 100-8901 Japan. Phone: (81-3) 3501-1361, Fax: (81-3) 3501-8391, e-mail: fujita-masahisa@rieti.go.jp

1 Introduction

> All theory depends on assumptions which are not quite true. That is what makes it theory. The art of successful theorizing is to make the inevitable simplifying assumptions in such a way that the final results are not very sensitive. A "crucial" assumption is one on which the assumptions do depend sensitively and it is important that the crucial assumptions be reasonably realistic. When the results of a theory seem to flow specifically from a special crucial assumption, then if the assumption is dubious, the results are suspect. (Solow, 1956, p. 65)

We wish to study how knowledge is created by researchers in a specific field of study. Such knowledge creation entails creative use of models and, as in the quote above, judicious use of assumptions. Although our primary examples come from the field of economic growth, we believe that our ideas are more widely applicable. The use of these examples helps to explain how economic research and models are similar to and different from their analogs in other research areas, such as the physical and natural sciences.

We are considering a pyramid of knowledge, as in Figure 1. Our primary focus is on the top level of the pyramid, working knowledge used in innovation. For example, this is the knowledge used by a researcher in the economics profession in developing new ideas. It relies on many more levels of foundational knowledge. *Background knowledge* includes, for example, basic math and arithmetic, that form part of the foundation for working knowledge. *Observational knowledge* contains, for example, data and information from the real world that might be related to a specific field of research. *General knowledge* includes basic knowledge needed to obtain a high school degree. *Social knowledge* includes language and culture.

The questions we seek to address are as follows. What is the appropriate way to formalize the *operational structure* of working knowledge space, the top of the pyramid, in which the *dynamic process* of innovation occurs? In this context, how are *specific* new ideas, research papers, and potential patents created by a research worker or a group of them, based on the current stock of knowledge? What roles do dynamics, heterogeneity of ideas, and heterogeneity of researchers play?

There are vast literatures related to the questions we ask. So we shall be selective in our literature review, with a focus on the differences between our work and the work of others. Broadly speaking, our paper might be seen as belonging to the general field of artificial intelligence or, more generally, the cognitive sciences. The triumph of classical artificial intelligence theory, initiated by Alan Turing and John von Neumann, was the invention of a machine that could reproduce itself. However, our perspective is very focused on how people, as opposed to machines, develop new ideas over time. In that sense, our work is also related to the humanities, in particular philosophy and the art of creation.

In talking about machines and people, as is well known, Turing (1936) showed the possibility of creating the so-called "universal Turing machine" that can imitate any computer. Building on Turing, von Neumann created the fundamental architecture for modern computers that can essentially perform the role of the universal Turing machine. If we consider the human brain to be a computer, then modern computers should be able to imitate it. But, in order to imitate the functioning of a human brain, its entire operation must be formally programmed in a way that can be read by machines. Yet, unfortunately, we know little about the formal process of knowledge creation by people.[1]

In this paper, with a focus on modern economic research as an example, we hope to initiate a small step in understanding the formal operational structure of the knowledge creation process. Research in modern economics is heavily dependent on formal models and mathematical analysis together with the econometric study of actual economies. Thus, on the one hand, in comparison with those fields in the humanities and the arts, modern economics is more suited to the formalization of innovation processes. On the other hand, relative to the literature in the natural sciences, surprisingly the concept of a "robot scientist" for biological research has been advocated by King et al. (2004) and Soldatoval et al. (2006). The work of Villaverde and Banga (2014) on reverse engineering in biology is closest to ours. But that kind of research model is primarily based on statistical correlations to determine underlying mechanisms in well-controlled environments. So research in economics fits between the extremes of the humanities and natural sciences, and thus represents a happy medium for analysis. Moreover, research in economics is highly

[1]See von Neumann (1958) for pioneering work on the comparison of the computer and the brain. The purpose of this paper, however, is not to compare the actual operation of a computer and that of the brain, but to express formally the process of knowledge creation by people so that someday a computer / brain can imitate it. See also Kurzweil (2005) for a stimulating presentation of recent developments in artificial intelligence.

dependent on context, in terms what is of economic importance at a particular time and location, as well as the personal characteristics of a researcher such as her knowledge stock, preferences and value judgements. We will be precise about this in the sequel. Thus, the concept of a "robot economist" might yet be premature.

The tools we employ come from many different areas. For instance, as in many applied areas, we use formal concept analysis, pioneered by Ganter et al (1998) and Ganter and Wille (1999). We note, however, that this tool tends to be static in nature, whereas our focus is on the dynamics of knowledge creation. Yet another aspect of our research is the use of Chu spaces; see, for example, the innovative work of Barwise and Seligman (1997) and Pratt (1999) for an explanation. As elaborated in section 3 below, we use four Chu spaces in sequence in our model. Our work is also closely related to philosophical and formal logic, specifically lattice theory and Boolean algebras, as elucidated in Watanabe (1969), Rott (2001) and Dunn and Hardegree (2001). Finally, our work is related to analytic function and singularity theory.

In order to address our questions, we introduce a *dynamic* framework for the fine microstructure of knowledge creation. This approach is a distinguishing feature of our work. First we identify a consistent knowledge state at a given time as a fixed point of a creation process. Then reverse engineering leads to the addition of new components to the knowledge creation process.

Our focus in this paper is only on positive theory. We relegate the important consideration of normative theory and the interrelation between normative and positive theory to future work.

The plan for the remainder of the paper is as follows. In the next section, we introduce our framework to formalize the operational structure of working knowledge space. In section 3, we begin our analysis of the model by deriving self-consistent knowledge states in the form of fixed points for a given person at a given time. In section 4, we prepare to analyze the dynamics of knowledge creation, introducing modifications to the static framework. Then in section 5, synthesizing the previous aspects of our framework, we examine the dynamic process of knowledge creation, allowing the knowledge structure to change in an endogenous way. In section 6, we discuss conclusions and possible extensions. For instance, we discuss the inclusion of normative as well as positive features of knowledge creation, as well as the extension to more than one researcher.

2 The Framework

2.1 A General Overview

The framework describing the dynamic process of innovation consists of several pieces. Figure 2 describes the macro picture (not yet with specifics) of our framework for the analysis of the dynamic process of knowledge creation. The framework consists of four quadrants in $\mathbb{R}^2$ reminiscent of the cobweb diagrams used in macroeconomics many years ago. We shall describe each quadrant in turn.

To begin, we consider the northwest quadrant, the one labelled "Theory." The horizontal axis $\mathcal{A}$ represents the set of all conceivable assumptions. For a person at any given time, the set of assumptions available to them will be a subset of $\mathcal{A}$. The vertical axis, given by $\mathcal{X}$, is the set of all possible models. Each element of $\mathcal{X}$ corresponds to a specific subset of $\mathcal{A}$, a collection of assumptions, that in turn is an element of the power set of $\mathcal{A}$. A *Theory* tells us which set of assumptions constitute which model.

Next consider the northeast quadrant, the one labelled "Science." We have already described the vertical axis $\mathcal{X}$, so next we shall describe the horizontal axis $\mathcal{Y}$. This axis represents the conceivable implications, namely the set of all possible conclusions that might be drawn from the solutions to any particular model. *Science* tells us how to obtain the implications of each model.

Moving to the southeast quadrant, the one labelled "Empirics," we have already described the horizontal axis $\mathcal{Y}$, so we next describe the vertical axis $\mathcal{Z}$. This axis represents available observations of the real world, either casual (as in stylized facts) or formal (as in data analysis). These observations might confirm or refute an implication. *Empirics* test the real world validity of the implications.

Finally, we discuss the southwest quadrant, labelled "Art." We have already discussed the axes $\mathcal{Z}$ and $\mathcal{A}$. *Art* allows us to select new assumptions based on observation. New assumptions are chosen step by step, slowly, as we modify known models to create new ones.

2.2 Preliminaries

This framework might seem very abstract and not terribly useful, so we shall now make it much more concrete using a real world example from the field of economic growth.

To begin, as stated previously, the elements of $\mathcal{A}$, $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ that are available to a person at a given time will not be the universe of all possibilities. Time will be discrete and indexed by $t = 1, 2, ...$ Agents will be indexed by i and j. We shall use $\mathcal{A}_i(t)$ to denote the set of assumptions available to person i at time t. Analogous notation applies for $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$.

We shall describe each of the four quadrants in turn, beginning with *Theory space*. Please refer to Figure 3, which corresponds to the early period of growth theory. Assumptions, namely elements of $\mathcal{A}_i(t)$, are denoted by $a_1, a_2, a_3 \in \mathcal{A}_i(t)$. Elements of $\mathcal{X}_i(t)$ are denoted by, for example, $x_1 = \{a_1, a_2\}, x_2 = \{a_1, a_3\} \in \mathcal{X}_i(t)$. Black dots reveal that a particular assumption, say a_1, is an element of a particular model, say $x_1 = \{a_1, a_2\}$. Dots with empty interior mean that a particular assumption, say a_2, is not an element of a particular model, say $x_2 = \{a_1, a_3\}$. Formally, these will be represented by 1 for a black dot and 0 for a dot with empty interior when we construct matrices to represent our various quadrants. There is an underlying mathematical structure that we shall describe in detail in the next subsection.

For illustration, we focus on our interpretation of the appearance of Solow (1956), one of the classics in economic growth.[2] To be precise, we shall focus on the *basic model* used in that paper, in contrast with the extensions. In order to make the intuition clear as we proceed through our model development, we shall be quite specific about this example. Under this interpretation, a_1 combines, for simplicity of expression, all basic assumptions common to early growth models, including a production function using both capital and labor, absence of technical change and exogenous population growth over time.[3] Assumption a_2 means a production function where the two factors, labor and capital, are used in fixed proportions. Assumption a_3 means a production function where the two factors are substitutable, for instance Cobb-Douglas, is used. Then the Harrod-Domar (H-D) model x_1 is represented by $\{a_1, a_2\} \in \mathcal{X}_S(t)$, where S stands for Solow and $t = 1955$. Eventually, Solow's model x_2 is represented by $\{a_1, a_3\} \in \mathcal{X}_S(1956)$.

There are two important points to make about the example represented in Figure 3. First, all models are assumed to share the basic assumptions a_1. Second, we have omitted singleton sets of assumptions $\{a_2\}$ and $\{a_3\}$ from models, since such models are tautological and hence not very useful to us. In other words, $\mathcal{X}_S(t)$ is a strict subset of all possible models.

[2] It is well-known that a similar paper was independently published by Swan (1956). We shall later attempt to explain why this occurred. See Dimand and Spencer (2008).

[3] We elaborate the basic assumptions in detail in section 4.1.

Next we turn to *Science space*, as illustrated in Figure 4. The horizontal axis represents the implications of the various models. Before getting into specifics, we must discuss precisely how to deal with those models (α) containing incompatible assumptions, and (β) having an incomplete specification (i.e., the set of assumptions of a model is not sufficient to define its solutions). Logically speaking, a model of type (α) or type (β) has no solution, so we put dots with empty interior in the entire row of Figure 4 associated with that model.

For example, the model $x_4 = \{a_1, a_2, a_3\}$ represented on the vertical axis has contradictory assumptions a_2 and a_3, so it has empty dots in its entire row. Likewise, the model $x_3 = \{a_1\}$ has an incomplete specification, so it has empty dots in its entire row. The Harrod-Domar model x_1 features "growing total GDP," "knife-edge path," and "expanding unemployment" away from the knife-edge growth path, so we put black dots in the matrix for these entries. Where conclusions do not apply, we put dots with empty interior. We shall explain how the Solow (1956) model fits into this later, when we discuss dynamics below.

The next quadrant is *Empirics*, as depicted in Figure 5. Here, the implications listed on the $\mathcal{Y}_i(t)$ axis are tested against the observations in the $\mathcal{Z}_i(t)$ axis. However, we do not test individual implications with individual observations. Rather, the set of implications of each model is tested against observations. The precise operation of this space is explained later in Section 3.

The final quadrant is *Art*, in the southwest, illustrated in Figure 6. This quadrant describes which assumptions are consistent with which observation and how the use of new assumptions arises from empirical observations. However, for testing the consistency between assumptions and observations, we need to employ "reverse engineering," looking back at the entire sequence of processes in Empirics space, Science space and Theory space in turn. The functioning of Art space is explained in detail in Section 3.

This completes our overview of the basic model in terms of an example. Before turning to analysis of the framework, we first describe its relationship to the mathematics literature.

2.3 The Underlying Mathematics

In preparation for the analysis of the model, we make a few remarks concerning the mathematical structure implied by our model.

If, in each quadrant, we replace each black dot by a 1 and each dot with empty interior with a 0, then each quadrant can be represented by a *Chu space*; see Barwise and Seligman (1997), Ganter et al. (1998), and Pratt (1999). Such spaces are common in the theory of computing, logic, and formal concept analysis. There have been a few applications to quantum mechanics and game theory.

In general, specific maps between Chu spaces are called *Chu transforms*. As Chu transforms and the general theory of Chu spaces is not yet used in the present analysis, we simply adopt the basic structure of a Chu space as applied to our context. It is quite simple, consisting of a matrix of zeros and ones for each quadrant.[4] We shall, however, use the important idea of a *formal concept* derived from a Chu space, that we shall introduce shortly.

We note that in general, there is no notion of distance or topology for elements of any axis.

2.4 Notation

Next we provide formal definitions for the general case. Let us start with the northwest quadrant, namely *Theory space*. We represent it generally by a Chu space $\mathbb{T} = \{\mathcal{A}, R_T, \mathcal{X}\}$, suppressing i and t for now for notational simplicity, where $\mathcal{A}$ represents the set of possible assumptions, $\mathcal{X}$ represents the set of possible models, and the binary relation R_T on $\mathcal{A} \times \mathcal{X}$ is interpreted as: aR_Tx means that assumption $a \in \mathcal{A}$ is included in the set of assumptions constituting model $x \in \mathcal{X}$. In general, we represent this by a matrix of 0's and 1's, where 1 means aR_Tx and 0 means *not* aR_Tx.

Following Ganter and Wille (1999), we define a formal concept derived from this Chu space as follows. A *formal concept* is a pair (A^*, X^*) with $A^* \subseteq \mathcal{A}$, $X^* \subseteq \mathcal{X}$ such that

$$\begin{aligned} A^* &= \{a \in \mathcal{A} \mid aR_Tx \text{ for all } x \in X^*\} \\ X^* &= \{x \in \mathcal{X} \mid aR_Tx \text{ for all } a \in A^*\} \end{aligned}$$

[4] In the course of learning about the elementary theory of Chu spaces and Chu transforms, the authors came across a number of important questions that do not seem to be addressed by the literature. First, given two arbitrary Chu spaces, does a Chu transform between them exist? The answer is apparently negative. Under what conditions do we know that a Chu transform does exist? Analogous to the theory of manifolds, are the properties of a Chu space independent of what larger space (e.g. more rows or columns that are all zero) it is embedded in?

The notion of a formal concept is very important to our analysis. The pair (A^*, X^*) represents a notion of duality. In general, A^* is the set of all objects in $\mathcal{A}$ that have all the attributes given in X^*, whereas X^* is the set of all attributes possessed by every object in A^*. Together, a formal concept (A^*, X^*) represents a unit of knowledge in this Chu space. For our first specific application in this paper, we interpret objects as assumptions, and attributes as models that might or might not contain these assumptions. For example, in Figure 3, $A^* = \{a_1, a_2\}$ and $X^* = \{x_1, x_4\}$ together are a formal concept. Formal concepts are not necessarily unique. However, the collection of all formal concepts forms a complete lattice itself (via set inclusion), and we shall use this important fact shortly. This lattice can be represented graphically by a Hasse diagram, visualizing cognitive structures in this Chu space.

Define $\mathcal{X}_0 \equiv \{\text{All models of type } (\alpha) \text{ or type } (\beta)\} \subseteq \mathcal{X}$, namely the set of all models with contradictory assumptions or incomplete assumptions. One of the main functions of *Theory space* is to divide $\mathcal{X}$, the set of all models, into $\mathcal{X}_0$ and its complement $\widehat{\mathcal{X}} = \mathcal{X} - \mathcal{X}_0$. The first set contains models that are not potentially solvable, whereas the second set contains models that are potentially solvable, hereafter called *viable models*, and thus of interest. In the context of Figure 3, for example, $x_3, x_4 \in \mathcal{X}_0$, whereas $\widehat{\mathcal{X}} = \{x_1, x_2\}$.

Analogous to *Theory space*, we can do something similar for the next quadrant, *Science space.* We represent it generally by a Chu space $\mathbb{S} = \{\mathcal{X}, R_S, \mathcal{Y}\}$, where $\mathcal{X}$ represents the set of possible models, $\mathcal{Y}$ represents the set of possible implications, and the binary relation R_S on $\mathcal{X} \times \mathcal{Y}$ where xR_Sy means that solutions to the model $x \in \mathcal{X}$ all have the implication $y \in \mathcal{Y}$. In general, we represent this by a matrix of 0's and 1's, where 1 means xR_Sy and 0 means *not* xR_Sy. A *formal concept* is a pair (X^*, Y^*) with $X^* \subseteq \mathcal{X}$, $Y^* \subseteq \mathcal{Y}$ such that

$$
\begin{aligned}
X^* &= \{x \in \mathcal{X} \mid xR_Sy \text{ for all } y \in Y^*\} \\
Y^* &= \{y \in \mathcal{Y} \mid xR_Sy \text{ for all } x \in X^*\}
\end{aligned}
$$

Before turning to a discussion of the remaining quadrants, we note a special feature of Science space. That is, as illustrated in Figure 4, the rows of models $x_3, x_4 \in \mathcal{X}_0$ consist entirely of 0's, meaning that *not* x_3R_sy and *not* x_4R_sy for all $y \in \mathcal{Y}$. Thus, models in $\mathcal{X}_0$ play no important role, and hence we focus on models in $\widehat{\mathcal{X}}$ in the following analysis.

The maps in *Empirics space* and *Art space* are very different from the first two maps.

In *Empirics space*, a map $f_{\mathbb{E}}$ takes subsets of $\mathcal{Y}$ to subsets of $\mathcal{Z}$. It maps sets of implications or conclusions to empirical observations that together are consistent with these conclusions. More precisely, the empirical observation z is in the image $f_{\mathbb{E}}(Y^*)$ exactly when it is consistent with the set of conclusions Y^*. This map is explained more precisely in the next section.

In *Art space*, a map $f_{\mathbb{A}}$ takes subsets of $\mathcal{Z}$ to subsets of $\mathcal{A}$. It maps empirical observations to assumptions that are implied or necessary conditions arising when the empirical observations are taken together. This map also is explained more precisely in the next section.

3 Static Knowledge States

To begin the analysis of our model, the first step is to define a self-consistent state of knowledge for a given person at a given time. Fortunately for us, our model yields a formal mathematical structure that corresponds well to the intuition. Mathematically, a *complete manuscript* is a fixed point of the sequence of four maps in the 4 quadrant diagram of Working-K space (Figure 2) given what is assumed to be known by a person at time t. Here we define the maps and their domains used by a person at time t. In the sections that follow, we shall allow these maps and domains to change with time. However, to keep notation simple, we shall often omit the time index except where this might cause confusion.

Let us explain the overall approach by using Figure 8. (Figure 7 is to be explained in the next section.) In this figure, the four diagrams in Figure 3 to Figure 6 are put together with some modifications. First, in *Theory space* in Figure 8, we put the diagram in Figure 3 after removing all rows associated with the models in $\mathcal{X}_0$. Thus, the vertical axis of *Theory space* (as well as of *Science space*) is now $\widehat{\mathcal{X}}$, representing all models that are potentially solvable. Second, after removing from Figure 4 all rows associated with the models in $\mathcal{X}_0$, we put Figure 4 in *Science space*, side by side with *Theory space* in Figure 8. Third, in *Empirics space* in Figure 8, we put the diagram in Figure 5 which is now filled with a matrix of black dots and empty dots. This matrix explains which set of implications (derived from *Science space*) are consistent with which set of conclusions. The derivation of this matrix, however, awaits the general formalization of *Empirics space* below. Likewise, in *Art space* in Figure 8, we put the diagram in Figure 6 which is now filled with a matrix of black dots and empty dots. This matrix associates each set of empirical observations (derived from *Empirics space*) with a set or sets of assumptions that are implied as necessary conditions arising when these observations are taken together. Again, however, the derivation of this matrix is to be explained in general terms below.

In general, we map the results of each quadrant to the next quadrant sequentially clockwise, and identify complete manuscripts as fixed points of the composition of these four maps. In deriving these maps, we fully utilize formal concept analysis in each quadrant.

We will begin with the northwest quadrant, namely *Theory space*. We shall again temporarily suppress i and t for simplicity.

Let $\mathbb{T} = \left\{\mathcal{A}, R_T, \widehat{\mathcal{X}}\right\}$ represent *Theory space*, in which the set of models is restricted to $\widehat{\mathcal{X}}$, representing all models that are potentially solvable. Let $\{(A_T^*, X_T^*)\}$ be the set of all formal concepts in *Theory space*, where each pair (A_T^*, X_T^*) with $A_T^* \subseteq \mathcal{A}$, $X_T^* \subseteq \widehat{\mathcal{X}}$ is such that

$$\begin{aligned} A_T^* &= \{a \in \mathcal{A} \mid aR_Tx \text{ for all } x \in X_T^*\} \\ X_T^* &= \left\{x \in \widehat{\mathcal{X}} \mid aR_Tx \text{ for all } a \in A_T^*\right\} \end{aligned}$$

Let $L_T^A = \{A_T^*\}$ be the collection of all such A_T^*, and $L_T^X = \{X_T^*\}$ the collection of all such X_T^*. Notice that each L_T^A and L_T^X forms a complete lattice under standard set-theoretic operations (Ganter and Wille, 1999).[5] For each pair $(A_T^*, X_T^*) \in L_T^A \times L_T^X$, A_T^* determines uniquely X_T^*, which is expressed as a map, $X_T^* = f_T(A_T^*)$, so that $f_T : L_T^A \to L_T^X$.

In the context of *Theory space* in Figure 8, we have the following three pairs of formal concepts:[6]

$$\begin{aligned} &A_{T_1}^* = \{a_1, a_2\}, \quad X_{T_1}^* = \{x_1\} = \{\text{HD}\} \\ &A_{T_2}^* = \{a_1, a_3\}, \quad X_{T_2}^* = \{x_2\} = \{\text{Solow56}\} \\ &A_{T_3}^* = \{a_1\}, \qquad X_{T_3}^* = \{x_1, x_2\} = \{\text{HD, Solow56}\} \end{aligned} \tag{1}$$

Next, let $\mathbb{S} = \left\{\widehat{\mathcal{X}}, R_S, \mathcal{Y}\right\}$ represent *Science space*. For each $X \subseteq \widehat{\mathcal{X}}$, define a pair $(X_{\mathbb{S}}^*(X), Y_{\mathbb{S}}^*(X))$ as follows:

$$\begin{aligned} Y_{\mathbb{S}}^*(X) &= \{y \in \mathcal{Y} \mid xR_Sy \text{ for all } x \in X\}, \\ X_{\mathbb{S}}^*(X) &= \{x \in \mathcal{X} \mid xR_Sy \text{ for all } y \in Y_{\mathbb{S}}^*(X)\}. \end{aligned}$$

[5] This means, for example, when $A_{T_1}^*, A_{T_2}^* \in L_T^A$, then $A_{T_1}^* \cup A_{T_2}^* \in L_T^A$ and $A_{T_1}^* \cap A_{T_2}^* \in L_T^A$.

[6] Formally speaking, $A_{T_4}^* = \{a_1, a_2, a_3, a_4\}$ and $X_{T_4}^* = \emptyset$ also forms a formal concept in *Theory space*. Keeping this fact in mind, however, we omit it in the following discussion for simplicity.

Then $(X^*_{\mathbb{S}}(X), Y^*_{\mathbb{S}}(X))$ is a formal concept. Define

$$f_{\mathbb{S}}(X) = Y^*_{\mathbb{S}}(X) = \{y \in \mathcal{Y} \mid xR_S y \text{ for all } x \in X\}$$

taking subsets of $\widehat{\mathcal{X}}$ to subsets of $\mathcal{Y}$. The map in *Science space* is analogous to the map in *Theory space*.

Let $L^X_S = \left\{X^*_{\mathbb{S}}(X) \mid X \in \widehat{\mathcal{X}}\right\}$ and $L^Y_S = \left\{Y^*_{\mathbb{S}}(X) \mid X \in \widehat{\mathcal{X}}\right\}$. Then, each L^X_S and L^Y_S forms a complete lattice. Let $f_S \circ f_T : L^A_T \to L^Y_S$ be the composition map defined by

$$f_S\left(f_T\left(A^*_T\right)\right) = Y^*_S\left(X^*_T\left(A^*_T\right)\right) \quad \text{for } A^*_T \in L^A_T,$$

which uniquely maps each set of assumptions in L^A_T to a set of implications in $\mathcal{Y}$.

In the context of Figure 8, we have that

$$\begin{aligned} f_S(f_T(A^*_{T_1})) &= f_S(\{x_1\}) = \{y_1, y_2, y_3\} \equiv Y^*_{S1} \\ f_S(f_T(A^*_{T_2})) &= f_S(\{x_2\}) = \{y_1, y_4, y_5\} \equiv Y^*_{S2} \\ f_S(f_T(A^*_{T_3})) &= f_S(\{x_1, x_2\}) = \{y_1\} \equiv Y^*_{S3} \end{aligned} \tag{2}$$

Next, *Empirics space* is represented by a Chu space $\mathbb{E} = \{L^Y_S, R_E, \mathcal{Z}\}$, which maps each set of implications in L^Y_S to a set of observations in $\mathcal{Z}$. That is, for each $Y^*_S \in L^Y_S$ and $z \in \mathcal{Z}$, $Y^*_S R_E z$ means that the empirical observation z is consistent with the set of implications Y^*_S. Here, however, we need to add an important note in constructing this Chu space, which is first explained in the context of Figure 8.

In the structure we have delineated on L^Y_S, we have not exploited the underlying set- and lattice- theoretic structure of the set of conclusions, since we have treated every subset of implications as a distinct and unrelated element of L^Y_S. Of course, these subsets can be related by set inclusion, intersection, and union. Here is an example. On the L^Y_S axis in Figure 8, we have three subsets of implications: $\{y_1, y_2, y_3\}$, $\{y_1, y_4, y_5\}$ and $\{y_1\}$. For the first two concepts, the correspondence between each concept and each individual observation on the $\mathcal{Z}$ axis is rather straightforward, based intuitively on "consistency with observation."[7] In contrast, $\{y_1\} = \{y_1, y_2, y_3\} \cap \{y_1, y_4, y_5\}$, meaning that $\{y_1\}$ is a "superconcept" of $\{y_1, y_2, y_3\}$ and $\{y_1, y_4, y_5\}$. Thus, in the $\{y_1\}$ column, for each observation, there is a black dot only when both subconcepts $\{y_1, y_2, y_3\}$

[7] For the precise formalization of this statement, see Section 4.3.

and $\{y_1, y_4, y_5\}$ support it. Generalizing this example, we assume that R_E in the *Empirics space* satisfies the following *consistency condition*:

> For all $Y_1^*, Y_2^*, Y_3^* \in L_S^Y$ and $z \in \mathcal{Z}$, when $Y_3^* = Y_1^* \cap Y_2^*$
> $Y_3^* R_E z$ if and only if $Y_1^* R_E z$ and $Y_2^* R_E^* z$

Now, let $\{(\mathcal{Y}_E^*, Z_E^*)\}$ be the set of all formal concepts in the Chu space E=$\{L_S^Y, R_E, \mathcal{Z}\}$, where each pair $(\mathcal{Y}_E^*, Z_E^*)$ with $\mathcal{Y}_E^* \subseteq L_S^Y$, $Z_E^* \subseteq \mathcal{Z}$ is such that

$$\begin{aligned}\mathcal{Y}_E^* &= \left\{Y \in L_S^Y \mid Y R_E z \text{ for all } z \in Z_E^*\right\} \\ Z_E^* &= \{z \in \mathcal{Z} \mid Y R_E z \text{ for all } Y \in \mathcal{Y}_E^*\}\end{aligned}$$

Let $L_E^Z = \{Z_E^*\}$ be the collection of all such Z_E^*, which forms a complete lattice. For each $Y_S^* \in L_S^Y$, define a pair $(\mathcal{Y}_E^*(Y_S^*), Z_E^*(Y_S^*))$ as follows:

$$\begin{aligned}Z_E^*(Y_S^*) &= \{z \in \mathcal{Z} \mid Y_S^* R_E z\} \\ \mathcal{Y}_E^*(Y_S^*) &= \left\{Y \in L_S^Y \mid Y R_E z \text{ for all } z \in Z_E^*(Y_S^*)\right\}\end{aligned}$$

Then $(\mathcal{Y}_E^*(Y_S^*), Z_E^*(Y_S^*))$ is a formal concept, and the set of all pairs $(\mathcal{Y}_E^*(Y_S^*), Z_E^*(Y_S^*))$ as Y_S^* varies over L_S^Y gives us all the formal concepts in *Empirics space.* We define the map $f_E : L_S^Y \to L_E^Z$ as follows:

$$f_E(Y_S^*) = Z_E^*(Y_S^*) = \{z \in \mathcal{Z} \mid Y_S^* R_E z\}$$

This defines the map in *Empirics space.*

In the context of Figure 8, we have that

$$\begin{aligned}f_E(\{y_1, y_2, y_3\}) &= \{z_1, z_2, z_3, z_4\} \equiv Z_{E1}^* \\ f_E(\{y_1, y_4, y_5\}) &= \{z_1, z_5, z_6, z_7\} \equiv Z_{E2}^* \\ f_E(\{y_1\}) &= \{z_1\} \equiv Z_{E3}^*\end{aligned} \tag{3}$$

So far, we have defined three maps:

The first map, $f_{\mathbb{T}} : L_T^A \to L_T^X$ maps from sets of assumptions to solvable models.

The second map, $f_{\mathbb{S}} : L_T^X \to L_S^Y$ takes a set of models to a set of implications that they all share.

The third map, $f_{\mathbb{E}} : L_S^Y \to L_E^Z$ takes each set of implications to their empirical observations.

Now, we can define the composition map $\Gamma \equiv f_E \circ f_S \circ f_T : L_T^A \to L_E^Z$ as follows:

$$\Gamma(A_T^*) = f_E(f_S(f_T(A_T^*))) \text{ for } A_T^* \in L_T^A.$$

When $\Gamma(A_T^*) \neq \emptyset$, the map Γ uniquely identifies the set of observations that are supported by the set of assumptions A_T^*. In contrast, when $\Gamma(A_T^*) = \emptyset$, it means that no observation is consistent with the set of assumptions A_T^*.

In the context of Figure 8, we have that[8]

$$\begin{aligned} \Gamma(\{a_1, a_2\}) &= \{z_1, z_2, z_3, z_4\} \equiv Z_{E1}^* \\ \Gamma(\{a_1, a_3\}) &= \{z_1, z_5, z_6, z_7\} \equiv Z_{E2}^* \\ \Gamma(\{a_1\}) &= \{z_1\} \equiv Z_{E3}^* \end{aligned} \tag{4}$$

The first equation above states that the set of conclusions $\{z_1, z_2, z_3, z_4\}$ can be supported together only by the Harrod-Domar model based on assumptions $\{a_1, a_2\}$. Likewise, the second equation says that the set of conclusions $\{z_1, z_5, z_6, z_7\}$ can be supported together only by the Solow 56 model based on assumptions $\{a_1, a_3\}$. In contrast, the last equation means that the observation z_1 can be supported by any solvable model (in this example, the Harrod-Domar model and the Solow 56 model) that share assumption a_1. This is because the basic set of assumptions includes the assumption of exogenous population growth over time.

Finally, provided that the range of map Γ is not empty, that is, $\Gamma(L_T^A) \equiv \cup_{A_T^* \in L_T^A} \Gamma(A_T^*) \neq \emptyset$, we wish to know which set of observations in L_E^Z can be supported by which set of assumptions in L_T^A. This means that we should conduct "reverse engineering" from L_E^Z to L_T^A, filling up the fourth quadrant. To do so, let f_T^{-1}, f_S^{-1} and f_E^{-1} be respectively the "inverses" of f_T, f_S and f_E.[9] Composing these inverse maps, let

$$\Gamma^{-1} \equiv f_T^{-1} \circ f_S^{-1} \circ f_E^{-1} : L_E^Z \to L_T^A$$

be the inverse of map Γ. Although Γ is a single-valued function, since each inverse map, f_T^{-1}, f_S^{-1} and f_E^{-1}, is a multi-valued map, in general, so is Γ^{-1}. This multi-valued map defines the *Chu space* $\mathbb{A} = \left\{L_E^Z, R_A, L_T^A\right\}$ in the fourth quadrant, where for $Z_E^* \in L_E^Z$ and $A_T^* \in L_T^A$,

$$Z_E^* R_A A_T^* \text{ when } A_T^* \in \Gamma^{-1}(Z_E^*).$$

In the context of Figure 8, we have that

$$\begin{aligned} \Gamma^{-1}(\{z_1, z_2, z_3, z_4\}) &= \{a_1, a_2\} \\ \Gamma^{-1}(\{z_1, z_5, z_6, z_7\}) &= \{a_1, a_3\} \\ \Gamma^{-1}(\{z_1\}) &= \{a_1\} \end{aligned} \tag{5}$$

[8] To be precise, we also have that $\Gamma(\{a_1, a_2, a_3, a_4\}) = \emptyset$.

[9] For example, $f_T^{-1}(X_T^*) = \left\{A_T^* \in L_T^A \mid f_T(A_T^*) = X_T^*\right\}$.

which are represented by black dots in the *Art space* of Figure 8.

Let F be the composition of two maps Γ and Γ^{-1}.

$$F = \Gamma^{-1} \circ \Gamma : L_T^A \to \mathcal{P}(L_T^A)$$

where $\mathcal{P}(L_T^A)$ denotes the power set of the set L_T^A. In general, F is a multi-valued map that takes each element of L_T^A to subsets of L_T^A. Note that F is, in general, not an identity map. For example, if $\Gamma(A_T^*) = \emptyset$ for $A_T^* \in L_T^A$, then $F(A_T^*) = \emptyset$.[10]

Let $\hat{L}_T^A = \left\{A_T^* \in L_T^A \mid \Gamma(A_T^*) \neq \emptyset\right\}$ be the domain of Γ. For any subset $S \subseteq \hat{L}_T^A$, define $F(S) = \cup_{A_T^* \in S} F(A_T^*)$. This extends the domain of F from $\hat{L}_T^A$ to $\mathcal{P}(\widehat{L}_T^A)$ and F can be considered to be a map from $\mathcal{P}(\widehat{L}_T^A)$ to $\mathcal{P}(\widehat{L}_T^A)$. The following always holds:

$$F(S^*) \supseteq S^*$$

If, for $S^* \subseteq \hat{L}_T^A$,

$$F(S^*) = S^*,$$

then we call S^* a *fixed point* of map F.[11] In this context, applying Berge (1963, Chapter II, Theorem 4), we can conclude as follows:

Theorem 1: Suppose that the domain of the map $\Gamma = f_E \circ f_S \circ f_T$ is not empty. Then, the fixed points of $F = \Gamma^{-1} \circ \Gamma$ form a complemented lattice.[12] *In particular, the set of fixed points is nonempty.*

This theorem implies that fixed points of F exist if and only if $\Gamma(L_T^A) \neq \emptyset$. In other words, fixed points of F exist if and only if the matrix of *Empirics space* contains at least one black dot, that is, at least one set of observations is supported by some set of assumptions.

In the context of Figure 8, combining (4) and (5), we can readily see that the map F has the following three atomic or elementary fixed points:[13]

$$\begin{aligned} S_1^* &\equiv \{a_1, a_2\} = A_{T_1}^* \\ S_2^* &\equiv \{a_1, a_3\} = A_{T_2}^* \\ S_3^* &\equiv \{a_1\} = A_{T_3}^* \end{aligned} \tag{6}$$

[10] We define $\Gamma^{-1}(\emptyset) = \emptyset$, and hence $F(A_T^*) = \emptyset$ when $\Gamma(A_T^*) = \emptyset$.

[11] In Berge (1963), such a set S is called a *stable subset*.

[12] A lattice is called "complemented" when: S^* is an element of a lattice implies that its complement is also an element of the lattice. Notice that since the number of the fixed points of F is finite, they form a complete lattice.

[13] To be precise, $S_4^* \equiv \emptyset$ is also a fixed point of F.

By definition, then, the following unions of these *atoms* of fixed points are also fixed points:

$$\begin{aligned} S_4^* &= \{A_{T_1}^*, A_{T_2}^*\},\ S_5^* = \{A_{T_1}^*, A_{T_3}^*\},\ S_6^* = \{A_{T_2}^*, A_{T_3}^*\}, \\ \text{and } S_7^* &= \{A_{T_1}^*, A_{T_2}^*, A_{T_3}^*\} = \hat{L}_T^A \end{aligned} \tag{7}$$

With the addition of $S_8^* = \emptyset$, all of these fixed points together form a complemented lattice. Given that the fixed points in (7) are derived from the atoms in (6), we may focus on the latter.

In the fourth quadrant of Figure 8, the three atomic fixed points are denoted by black dots. The fixed point $S_1^* \equiv \{a_1, a_2\}$ corresponds to the Harrod-Domar model, and it uniquely supports the set of implications $\{z_1, z_2, z_3, z_4\}$. Likewise, $S_2^* \equiv \{a_1, a_3\}$ corresponds to the Solow 56 model, supporting uniquely the set of implications $\{z_1, z_5, z_6, z_7\}$. In contrast, the fact that the fixed point $S_3^* \equiv \{a_1\}$ supports z_1 indicates that the observation z_1 can be supported by any solvable model, including the Harrod-Domar model and the Solow 56 model, in which assumption a_1 is used. Recall that a_1 represents the basic set of assumptions including the assumption of growing population.

Why are we interested in fixed points, in particular elementary or atomic fixed points? It is because a fixed point is internally consistent, and is taken as far as possible given a person's knowledge at that time. By internal consistency we mean that the assumptions, model, implications, and empirical observations are all consistent with each other. There is no immediate cause for alteration.

We also note that, since the domain of F is L_T^A, and the set of fixed points is a complemented lattice, the union of all the elements of all the fixed points covers $\mathcal{A}$.

As illustrated above, fixed point analysis utilizing formal concepts in the four quadrants of working knowledge space reveals well the structure of knowledge possessed by a researcher at a given time.

4 Some Preparations for Dynamics

Up to this point we have taken a person's state of knowledge at time t to be fixed. That is, the matrices representing the Chu space in each quadrant do not change. Next, we intend to consider dynamics, allowing the matrices to change endogenously. Jumping from statics to dynamics, however, is not easy. As an intermediate step, in this section, we introduce several concepts and tools that are useful in understanding the dynamic process of knowledge creation.

In the first subsection, we introduce Boolean algebras in Theory space, and consider how one can systematically introduce new assumptions sequentially. In the second subsection, we introduce Boolean algebras in Science space, and identify potential sets of implications that could be derived from possible new models. In the third subsection, we introduce utility functions in Empirics space, and consider how a researcher would choose new specific assumptions or models. In the last subsection, we consider the interactions between the 4 quadrant working space and supporting knowledge spaces.

4.1 Complementation in $\mathcal{A}$ and Boolean algebras in Theory space

When a researcher introduces new assumptions and develops new models sequentially, one may wonder where these new assumptions might come from. Our understanding is that new assumptions do not need to come from an extraterrestrial; actually, they are already hidden behind the existing set of assumptions. To elaborate this point, recall that in the context of Figure 3 in Section 2.2, we stated that a_1 represents *the set of all basic assumptions* that are common to early growth models. To be precise, in the context of Solow (1956), the following set of assumptions are explicitly stated in Section II of that paper:

one consumption commodity, constant savings ratio, two factors of production: capital and labor, production using constant returns to scale in net output, absence of technological change, labor force increasing at a constant relative rate, inelastic supply of labor, efficient utilization of production factors[14]*, competitive pricing of factors.*

Formally speaking, however, Solow was assuming many more basic assumptions that he did not, understandably, bother to mention in Section II. These *implicit basic assumptions* include:

each factor is homogenous, absence of technological externality, closed economy, no space, no transport costs, continuous time, all variables are measured

[14] In Solow (1956) Section IV, the Harrod-Domar case of fixed factor proportions in production is introduced as the first example: $Y = F(K, L) = \min\left(\frac{K}{a}, \frac{L}{b}\right)$. In this case, except when $\frac{K}{L} = \frac{a}{b}$, the two factors cannot be fully employed. Therefore, when this Harrod-Domar case is permitted in Section II, full employment of both factors cannot be a part of the basic assumptions. Thus, we replace it here by the more general assumption of "efficient utilization of production factors," meaning that the maximum rate of output using the given level of factors is achieved.

in continuous numbers[15]

In general, let

$$a_1 = \{a_{11}, a_{12}, \cdots, a_{1i}, \cdots, a_{1m}\}$$

represent the set of all basic assumptions, including both explicit and implicit ones. Logically speaking, when one sets a_1 above, one automatically sets its *negation*,

$$\overrightharpoon{a}_1 = \left\{\overrightharpoon{a}_{11}, \overrightharpoon{a}_{12}, \cdots, \overrightharpoon{a}_{1i}, \cdots, \overrightharpoon{a}_{1m}\right\}$$

where

$$\overrightharpoon{a}_{1i} = \text{the negation of } a_{1i},\ i = 1, 2, \cdots m$$

Here, given an assumption a_{1i}, it is important to distinguish between its negation $\overrightharpoon{a}_{1i}$ and its *contraries.*[16] For example, let a_{1i} be the basic assumption:

$$a_{1i} = \text{absence of technological change}$$

Then, there are many (potentially infinitely many) contraries of a_{1i} that are incompatible with a_{1i}. For example, contraries of a_{1i} (absence of technological change) include any form of exogenous technological change such as neutral technological change, labor-augmenting technological change, capital-augmenting technological change, any form of embodied technological change, as well as any form of endogenous technological change. Therefore, *the negation of a_{1i}, the weakest proposition inconsistent with a_{1i}, is the disjunction of all of its contraries*:

$$\begin{aligned}\overrightharpoon{a}_{1i} &= \text{presence of some form of technological change} \\ &= \{a_{1i}^{c_1}, a_{1i}^{c_2}, \cdots, a_{1i}^{c_k}, \cdots\}\end{aligned}$$

where each $a_{1i}^{c_k}$ represents a specific form of technological change. In general, the negation of proposition a_{1i} contains a large number (possibly an infinite number) of contraries.

When one sets the basic assumption, a_1, its negation $\overrightharpoon{a}_1$ is to be registered automatically in the *stock of knowledge* associated with Theory space.[17] We could introduce any component of $\overrightharpoon{a}_1$, say $\overrightharpoon{a}_{1i}$, as a member of $\mathcal{A}$ (the set of all assumptions conceived explicitly). However, in general, $\overrightharpoon{a}_{1i}$ contains many

[15] In other words, there are no indivisibilities.

[16] This part of our discussion is based on Dunn and Hardegree (2001, p. 89).

[17] We elaborate on the stock of knowledge associated with each subspace of Figure 2 in Section 4.4.

contraries of a_{1i}, and hence $\overrightarrow{a}_{1i}$ itself is not specific enough to be a component of solvable models. Hence, when one introduces into $\mathcal{A}$ an alternative to a_{1i}, a specific component of $\overrightarrow{a}_{1i}$ must be chosen.

Actually, what was said above holds not only for a_1, but also for any existing member a_i of $\mathcal{A}$.[18] That is, for each $a_i \in \mathcal{A}$, its negation $\overrightarrow{a}_i$ is also to be registered as a part of the stock of knowledge associated with Theory space. Therefore, given the set of all possible assumptions,

$$\mathcal{A} = \{a_1, a_2, \cdots, a_n\},$$

its entire negation,

$$\overrightarrow{\mathcal{A}} = \left\{\overrightarrow{a}_1, \overrightarrow{a}_2, \cdots, \overrightarrow{a}_n\right\},$$

must be registered as a part of the stock of knowledge associated with Theory space.

If we set

$$\mathcal{A}^u = \mathcal{A} \cup \overrightarrow{\mathcal{A}}$$

then it holds by definition that

$$\mathcal{A} \cap \overrightarrow{\mathcal{A}} = \emptyset \text{ and } \overrightarrow{\mathcal{A}} \cup \mathcal{A} = \mathcal{A}^u$$

Thus, $\mathcal{A}$ and $\overrightarrow{\mathcal{A}}$ together form a partition of *the universe* of conceivable assumptions.

One might then wonder why we do not use $\mathcal{A}^u$, instead of $\mathcal{A}$, as the set of all conceivable assumptions from the start. If we did so, then since $\mathcal{A}^u$ seems to represent the universe of all conceivable assumptions, we would not need to worry about introducing new assumptions or models anymore. Although this is a legitimate idea, there are a few difficulties both operationally and conceptually.

First, as illustrated in the case of a_1, in general, each $\overrightarrow{a}_i$ contains many (possibly infinitely many) contraries of a_i. Therefore, $\overrightarrow{\mathcal{A}}$ contains quite a large number (possibly an infinite number) of elements. Thus, practically speaking, it is not possible to study all possible combinations of the elements of $\mathcal{A}^u$ at the same time. Indeed, in practice, finding even one potentially meaningful new set of assumptions and analyzing the associated model might involve a great deal of effort for a researcher. Second, in general, given $a_i \in$

[18] We are implicitly assuming that a_1 represents the set of basic assumptions that are mutually compatible, whereas each a_i $(i \geq 2)$ consists of a single proposition. However, it does not need to be so. Some of the a_i $(i \geq 2)$ may also consist of multiple propositions / assumptions that are mutually compatible. It is a matter of convenience in expression.

$\mathcal{A}$, specifying concretely the content of its negation $\overrightarrow{a}_i$ in terms of contraries requires a thorough knowledge of the current state of literature, whereas the state of the literature will change with time. For example, suppose that

$$a_{1i} = \text{perfectly competitive markets for all factors and goods}$$

Then, one can write down automatically that

$$\overrightarrow{a}_{1i} = \text{some markets are not perfectly competitive}$$

However, in the 1950's, specifying even one operational form of a noncompetitive market in the context of growth theory was not easy. This became possible only after sufficient progress in industrial organization theory and related fields in the 1960's and 1970's. In general, therefore, the content of $\mathcal{A}^u$ depends on the researcher's general knowledge as well as on the state of literature in related fields, and hence it evolves with time.

When a researcher specifies the set $\mathcal{A}$, she may register $\overrightarrow{\mathcal{A}}$ as concretely as possible as a part of the knowledge stock associated with Theory space, and renew $\overrightarrow{\mathcal{A}}$ with time. Then, transferring specific members of $\overrightarrow{\mathcal{A}}$ to $\mathcal{A}$ sequentially, the content of $\mathcal{A}$ can be enriched endogenously almost endlessly.

That said, however, it may not be easy to select meaningful combinations of the elements of $\mathcal{A}$ as potentially solvable models. In order to achieve a systematic transfer of members of $\overrightarrow{\mathcal{A}}$ to $\mathcal{A}$ and to choose meaningful combinations, it is useful to consider the process in terms of Boolean lattices or Boolean algebras. These constructs are closely related to the formal concept analysis presented in the previous section.

Recall that a lattice L with join $\vee$ and meet $\wedge$ is called a Boolean lattice or *Boolean algebra* if

(i) L is distributive (i.e., for any $a, b, c \in L, a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$)

(ii) L has the top element, 1, and the bottom element, 0.

(iii) Each $a \in L$ has a unique *complement* $\tilde{a}$ such that $a \vee \tilde{a} = 1$ and $a \wedge \tilde{a} = 0$.

For our purposes, an important characteristic of a Boolean algebra is that *any closed subinterval is also a Boolean algebra.* That is, for any $a, b \in L$ such that $a \leq b$, the closed interval,

$$[a, b] = \{a \leq x \leq b\},$$

also constitutes a Boolean algebra where b now becomes the top element, and a the bottom element. It is also well-known that for any set with a finite number of elements, its power set $\mathcal{P}(A)$ is a Boolean algebra with join and meet defined by set-union and set-intersection, respectively.

Now, going back to the early stages of economic growth theory development, recall that the combination of the next two assumptions,

$$\begin{aligned} a_1 &= \text{the set of basic assumptions} \\ a_2 &= \text{fixed factor ratio production technology,} \end{aligned}$$

defines the Harrod-Domar model. Here, a_1 includes all explicit and implicit assumptions, explained earlier, that are stated in Section II of Solow (1956). In contrast, a_2 represents a specification or specialization of the assumption of constant-returns-to-scale production technology included in a_1. The two assumptions a_1 and a_2 together form a solvable growth model.

Next, let us consider enriching the simple set of possible assumptions, $\mathcal{A} = \{a_1, a_2\}$, by transferring $\overline{a}_2$ from $\overline{\mathcal{A}}$, the stock of knowledge associated with Theory space, and adding it to $\mathcal{A}$. Strictly speaking, however, the negation of a_2,

$$\overline{a}_2 = \text{substitutable production technology}$$

is not specific enough to be a part of a solvable model. In fact, as a contrary of a_2, Solow (1956) considered what is now called the neoclassical production technology that can be expressed by a sufficiently smooth concave production function:[19]

$$a_2^c = \text{neoclassical production technology}$$

Then, we have new $\mathcal{A} = \{a_1, a_2, a_2^c\}$ such that

$$\begin{aligned} a_1 &= \text{the set of basic assumptions} \\ a_2 &= \text{fixed-factor-ratio production technology} \\ a_2^c &= \text{neoclassical production technology.} \end{aligned}$$

where a_2^c is a contrary of a_2.

Figure 9(a) depicts the Hasse diagram of the power set, $\mathcal{P}(\{a_1, a_2, a_2^c\})$, known as the *cube*.

[19] To be precise, in Solow (1956), many different forms of a substitutable production technology are considered. But here let us focus on the so called neoclassical production technology defined, for example, by Burmeister and Dobell (1970). Given the input vector $x = (x_1, x_2, \cdots, x_i, \cdots, x_n)$, the production function $F(x)$ is assumed to be linearly homogenous and concave with $F(0) = 0$, with continuous partial derivatives, and with strictly convex isoquants.

Figure 9

We can readily check that this cube represents a Boolean algebra. However, the entire cube does *not* form a Boolean algebra *for a potentially solvable model* because its top element, $\{a_1, a_2, a_2^c\}$, contains contradictory assumptions, specifically a_2 and a_2^c. In this cube, we can identify two closed intervals of Boolean algebras, each satisfying the following three conditions:

Condition 1: it contains no pair of incompatible assumptions.
Condition 2: it contains *the set of basic assumptions* at the bottom.
Condition 3: it represents a *maximal* closed interval that satisfies Condition 1 and Condition 2.

When a closed interval or *Boolean sub-algebra* of $\mathcal{P}(\mathcal{A})$ satisfies the three conditions above, as before we call it a *viable model.* The necessity of Condition 1 for a viable model is obvious. Condition 2 means that any viable model should contain the set of basic assumptions for it to be a solvable complete model. Condition 3 means that there is no larger Boolean sub-algebra in $\mathcal{P}(A)$ that satisfies Conditions 1 and 2. In adopting these three conditions, we are *playing it safe* in the sense that if a model contains redundant assumptions, then they will be identified later in the actual analysis.[20]

In Figure 9(a), Boolean sub-algebras of two viable models are represented respectively by a bold line:

Harrod Domar model	:	$\{\{a_1\}, \{a_1, a_2\}\}$
Solow 56 model	:	$\{\{a_1\}, \{a_1, a_2^c\}\}$

In Figure 9, diagram (b) represents formal concept analysis that has been transferred from Theory space in Figure 8. In comparing the two diagrams (a) and (b), we can see that each black dot in (a) corresponds to a formal concept (i.e., an element of L_T^A) in diagram (b). This is not surprising, for Boolean sub-algebras of viable models in (a) represent the Hasse diagram of formal concepts in (b).[21] Mathematically speaking, therefore, the study of Theory space in terms of Boolean algebras is identical with that in terms of formal concepts. Diagram (a) provides more information on where viable models are

[20] This point is elaborated later in Examples 1 and 2.

[21] Strictly speaking, for L_T^A to represent a complete lattice of concepts, it should also include $\{a_1, a_2, a_2^c\}$. Likewise, if the top element $\{a_1, a_2, a_2^c\}$ were added to the three black dots in diagram (a), then the four dots together would form a larger Boolean algebra. However, in both diagrams, $\{a_1, a_2, a_2^c\}$ does not correspond to a viable model.

situated in the entire Boolean algebra of $\mathcal{P}(\mathcal{A})$. However, as we will see next, very soon the entire graph of $\mathcal{P}(\mathcal{A})$ will become too big to handle when the number of elements of $\mathcal{A}$ increases.

In the preceding example, we enriched the set $\mathcal{A}$ by transferring a contrary, a_2^c, of assumption a_2 from $\overline{\mathcal{A}}$ to $\mathcal{A}$. Let us consider next the transfer of a contrary of a component of a_1 (= the set of basic assumptions), say a contrary of a_{11}, to enlarge $\mathcal{A}$. For example, let

$$a_{11} = \text{absence of technological change}$$

As noted before, its negation, $\overline{a}_{11}$, contains many contraries, and hence we must choose a specific one. Following Solow (1957), as a contrary of a_{11}, let us introduce

$$a_{11}^c = \text{the original production function shifts upward at a constant rate}$$

Since a_{11} and a_{11}^c are incompatible, the original a_1 and a_{11}^c are incompatible. To make the set of basic assumptions compatible with a_{11}^c, we separate a_{11} from a_1 and define $a_1' = a_1 - a_{11}$ to be the new set of basic assumptions. Then, the new set of all possible assumptions is now

$$\mathcal{A} = \{a_1', a_{11}, a_{11}^c, a_2, a_2^c\}$$

where

$$\begin{aligned} a_1' &= a_1 - a_{11} = \text{the new set of basic assumptions} \\ a_{11} &= \text{absence of technological change} \\ a_{11}^c &= \text{the original production function shifts upward at a constant rate} \\ a_2 &= \text{fixed factor ratio production technology} \\ a_2^c &= \text{neoclassical production technology} \end{aligned}$$

In this context, diagram (a) of Figure 10 depicts the Boolean algebra of the entire power set $\mathcal{P}(\mathcal{A})$. Within this entire power set, bold lines delineate four Boolean sub-algebras that meet the three conditions of viable models:

Harrod-Domar model	:	$\{\{a_1'\}, \{a_1', a_{11}\}, \{a_1', a_2\}, \{a_1', a_{11}, a_2\}\}$
Solow 56 model	:	$\{\{a_1'\}, \{a_1', a_{11}\}, \{a_1', a_2^c\}, \{a_1', a_{11}, a_2^c\}\}$
HD′ model	:	$\{\{a_1'\}, \{a_1', a_2\}, \{a_1', a_{11}^c\}, \{a_1', a_{11}^c, a_2\}\}$
Solow 57 model	:	$\{\{a_1'\}, \{a_1', a_2^c\}, \{a_1', a_{11}^c\}, \{a_1', a_{11}^c, a_2^c\}\}$

Figure 10

In Figure 10, diagram (b) represents the formal concepts in the Chu space of four viable models. As in Figure 9, each black dot in diagram (a) corresponds to a formal concept in diagram (b), an element of L_T^A.[22]

Thus far, in deriving viable models from a given set of assumptions, $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$, we did not pay much attention to the logical relations among assumptions. To be precise, we first considered the Boolean algebra of $\mathcal{P}(\mathcal{A})$ based purely on the set-theoretic combinations of assumptions. Then, we selected viable models as Boolean sub-algebras of $\mathcal{P}(\mathcal{A})$ that satisfy Conditions 1, 2, and 3. Among the three conditions, only Condition 1 reflects a logical relation, i.e., incompatibility, among assumptions. In general, however, given a pair of assumptions, a_i and a_j ($a_i \neq a_j$), when they are logically related, we can consider the following three possible cases:[23]

incompatible	: $a_i \wedge a_j = \emptyset$
specification	: $a_i \wedge a_j = a_i$ or $a_i \wedge a_j = a_j$ (i.e., $a_i \rightarrow a_j$ or $a_j \rightarrow a_i$)
partly compatible	: $a_i \wedge a_j \neq \emptyset, a_i \wedge a_j \neq a_i, a_i \wedge a_j \neq a_j$

The first case of incompatibility has been incorporated in Condition 1. To consider the remaining two cases, let us examine two examples below in which we add a new assumption a_4 to the set of assumptions, $\mathcal{A} = \{a_1, a_2, a_2^c\}$, that has been discussed in the context of Figure 9.

Example 1: $\mathcal{A} = \{a_1, a_2, a_2^c, a_4\}$ where

$$\begin{aligned} a_1 = &\ \text{the set of basic assumptions} \\ a_2 = &\ \text{fixed-factor-ratio production technology} \\ a_3 = a_2^c = &\ \text{neoclassical production technology} \\ a_4 = a_2^{c'} = &\ \text{Cobb-Douglas production technology} \end{aligned}$$

Clearly, a_4 is a specification of a_3 (i.e., $a_4 \rightarrow a_3$). Figure 11 depicts the Boolean algebra of the power set $\mathcal{P}(\mathcal{A})$, where the Boolean sub-algebras satisfying Conditions 1, 2 and 3 are delineated by bold lines. Among the two sub-algebras, the one involving a bold, broken line needs special attention. In the present context, since $a_4 \rightarrow a_3$ (i.e., $a_3 \wedge a_4 = a_4$), the model defined by $\{a_1, a_3, a_4\}$

[22] Again, formally speaking, L_T^A should contain $\mathcal{A}$ as an element. Likewise, in diagram (a), when $\mathcal{A}$ is added to the elements connected by bold lines as the common top element, then the entire graph represents a complete lattice. It is not, however, a Boolean algebra because it is not a closed interval of $\mathcal{P}(\mathcal{A})$.

[23] Here, the conjunction ($\wedge$) should be taken in an appropriate set of elements that are partially ordered.

is identical to the model $\{a_1, a_4\}$. Therefore, the sub-algebra with the top element $\{a_1, a_3, a_4\}$ can be decomposed into two Boolean sub-algebras, defined by the closed intervals, $[\{a_1\}, \{a_1, a_3\}]$ and $[\{a_1\}, \{a_1, a_4\}]$, respectively. Thus, as represented in Figure 11 (b-1), there exists three viable models. Although x_3 is a specification of x_2, the two models are different. In fact, x_2 represents the Solow 56 model, whereas x_3 corresponds to the model of Swan (1956).
Example 2: $\mathcal{A} = \{a_1, a_2, a_2^c, a_4\}$ where

$$\begin{aligned}
a_1 &= \text{ the set of basic assumptions} \\
a_2 &= \text{ fixed-factor-ratio production technology} \\
a_3 = a_2^c &= \text{ neoclassical production technology} \\
a_4 = a_2^{c''} &= \text{ CES production technology such that} \\
& \quad F(K, L) = \left[\alpha K^\beta + (1-\alpha) L^\beta\right]^{1/\beta}, 0 < \alpha < 1, 0 < \beta \leq 1
\end{aligned}$$

where K represents capital usage and L represents labor usage. In this context, $a_4 \to a_3$ when $\beta < 1$. However, when $\beta = 1$, $F(K, L) = \alpha K + (1-\alpha)L$, which does not belong to the neoclassical production technology. Therefore, a_3 and a_4 are partly compatible with each other. Figure 11(a) represents the Boolean algebra of the power set $\mathcal{P}(\mathcal{A})$, which seems identical to the one in the context of Example 1. However, in the present context, $a_3 \wedge a_4 \neq a_3$ and $a_3 \wedge a_4 \neq a_4$. Therefore, the Boolean sub-algebra with the top element $\{a_1, a_3, a_4\}$ can be decomposed into three different closed intervals, $[\{a_1\}, \{a_1, a_3\}]$, $[\{a_1\}, \{a_1, a_4\}]$ and $[\{a_1\}, \{a_1, a_3, a_4\}]$, each one representing a different viable model as shown in Figure 11(b-2).

In summary of the discussion above, given a set of assumptions $\mathcal{A}$, we may identify the set of viable models in two steps:

Step 1: Identify viable models as closed Boolean sub-algebras of $\mathcal{P}(\mathcal{A})$ that satisfy Conditions 1, 2 and 3.

Step 2: Considering the logical relations of specification or partial compatibility between each pair of assumptions, decompose each viable model into a set of independent models.

Finally, in forming new models by adding assumptions sequentially, we may differentiate "small changes" from "big changes." A small change in a model (or a variation) maintains the "core assumptions" in the basic set of assumptions, and introduces a variation of a "non-core" assumption. In contrast, a big change in a model replaces some of the "core assumptions" contained in the basic set of assumptions. In growth theory, for example, replacing the assumption of constant returns to scale with increasing returns to scale will lead to a big change in the model. In terms of Kuhn (1962), the former may represent "normal science," whereas the latter may represent a "paradigm shift."

4.2 Boolean algebras in Science space

Suppose that we have fixed the set of viable models, $\mathcal{X}$, from Theory space. Then, the next job is to examine the implications of each model in Science space. If the set of implications, $\mathcal{Y}$, had been prepared beforehand, this job would be easier. The question, however, arises: How has the set $\mathcal{Y}$ been prepared, and how complete is the set? To be prepared for all possible implications from all possible models, the set would need to be quite large. Furthermore, a new model might yield new implications that are not listed initially in $\mathcal{Y}$.

In conducting analysis in Science space, we may consider another approach. Instead of preparing the set $\mathcal{Y}$ beforehand, we may create $\mathcal{Y}$ more or less from scratch. That is, given a viable model, say $x_1 \in \mathcal{X}$, we obtain the solution of x_1, and make a list of its major implications. We do this work for each $x_i \in \mathcal{X}$ in turn. Then, putting all implications together, we get $\mathcal{Y}$.

The latter approach might sound easier. However, it too is not easy in practice. That is, without having some a priori list of implications to be investigated, a researcher would be unable to find the major implications of each model. Where should the researcher look?

In practice, a researcher would take a middle road, combining the two approaches. When conducting research, she is working in a specific field in which some stock of knowledge has been accumulated already. Furthermore, when working in Science space, she examines the implications of one specific new model at a time, whereas the major implications of the balance of existing models are commonly known among researchers in the field. Thus, we can expect that when a researcher starts examining implications of a new model, a preliminary set of possible implications, say $\mathcal{Y}(t-1)$, is already available. Through the actual examination of the implications of that model, then, she may find new implications that are to be added to $\mathcal{Y}(t-1)$, forming $\mathcal{Y}(t)$.

In this process of constructing the set $\mathcal{Y} \equiv \mathcal{Y}(t)$, it is important to consider logical relations among the elements of $\mathcal{Y}$. This consideration is also useful for finding potential sets of implications that would be possessed by new possible models.

For example, let us go back to an early stage of the evolution of economic growth theory, as depicted in Figure 7. Phase 1 is dated 1955. There exists only one model, x_1, the Harrod-Domar model, on which Robert Solow focused. It was agreed in the literature that important implications of H-D models were:

$$\begin{aligned} y_1 &= \text{growing total GDP} \\ y_2 &= \text{knife-edge growth path} \\ y_3 &= \text{expanding unemployment of a factor} \end{aligned}$$

Logically speaking, when these three implications of the H-D model are listed, one can automatically consider their *negations*:

$$\begin{aligned} \overline{y}_1 &= \text{not growing total GDP} \\ \overline{y}_2 &= \text{not knife-edge growth path} \\ \overline{y}_3 &= \text{expanding unemployment of no factor} \end{aligned}$$

These negations may be included in $\mathcal{Y}$ or recorded in the stock of knowledge associated with Science space. However, one may not include $\overline{y}_1$ in $\mathcal{Y}$ because any growth model should imply y_1, not $\overline{y}_1$; $\mathcal{Y}$ should be kept simple without unnecessary clutter. One may include $\overline{y}_2$ and $\overline{y}_3$ in $\mathcal{Y}$. However, $\overline{y}_2$ is too general, containing all *contraries* of y_2. The same is true for $\overline{y}_3$. Thus, we may imagine that Robert Solow selected y_2^c and y_3^c below as specifications of contraries to y_2 and y_3, respectively:[24]

$$\begin{aligned} y_2^c &= \text{converging to balanced-growth path} \\ y_3^c &= \text{full employment of both factors} \end{aligned}$$

We may imagine that Solow selected these specific contraries because he hoped that his new model had these implications, as we shall explain further in the next subsection.

We then have the set of implications as follows:

$$\mathcal{Y} = \{y_1, y_2, y_3, y_2^c, y_3^c\}$$

Since $\mathcal{Y}$ consists of five elements, the Boolean algebra of the power set, $\mathcal{P}(\mathcal{Y})$, is isomorphic to that of $\mathcal{P}(\mathcal{A})$ in Figure 10 (a). (Please neglect the bold lines in the diagram.) In this Boolean algebra of $\mathcal{P}(\mathcal{Y})$, as before, we select closed intervals, each satisfying the following three conditions:

Condition 1:	it contains no pair of incompatible implications.
Condition 2:	it contains y_1 (the basic implication) at the bottom.
Condition 3:	it represents a maximal closed-interval that satisfies Condition 1 and Condition 2.

[24] Notice that constant unemployment of a factor is also a contrary of y_3.

When a closed interval or Boolean sub-algebra of $\mathcal{P}(\mathcal{Y})$ satisfies the three conditions above, we may call it the set of conceivable implications or a *viable implication-set*.

In the present context of $\mathcal{Y} = \{y_1, y_2, y_3, y_2^c, y_3^c\}$, y_2 and y_2^c as well as y_3 and y_3^c are mutually incompatible. Furthermore, y_3 and y_2^c are also incompatible. Therefore, as depicted in Figure 12 (a), we can select three viable implication-sets from $\mathcal{P}(\mathcal{Y})$; they correspond to those delineated by bold lines in Figure 10 (a). The top element $\{y_1, y_2, y_3\}$ represents the actual implications of Harrod-Domar model. The top element $\{y_1, y_2^c, y_3^c\}$ represents the viable implication-set of a potential model x_2° about which we do not yet know. Similarly, the top element $\{y_1, y_2, y_3^c\}$ represents the viable implication-set of another potential model x_3°. Figure 12 (b) shows these viable implication-sets in Science space. The potential models x_2° and / or x_3° must be examined later through reverse-engineering.

In the preceding example, by using Figure 12 we extended the original set of implications, $\{y_1, y_2, y_3\}$ of the Harrod-Domar model, to $\{y_1, y_2, y_3, y_2^c, y_3^c\}$ by introducing contraries, y_2^c and y_3^c. One can also extend the set of implications by introducing specifications or a partition of an element, say y_i. For example, in the present context of the Harrod-Domar model, it is natural to consider the following partition or specifications of the implication y_1:

$$y_{11} = \text{growing total GDP but not per capita}$$
$$y_{12} = \text{growing total and per capita GDP}$$

where $y_{11} \wedge y_{12} = \emptyset$ and $y_{11} \vee y_{12} = y_1$. In this case, the set of implications is expanded to $\mathcal{Y}^+ = \{y_1, y_2, y_3, y_2^c, y_3^c, y_{11}, y_{12}\}$. In this context, each of the rows in Figure 12(b) can be extended by adding either y_{11} or y_{12}. Thus, as depicted in Figure 13, we obtain 6 rows or potential models: x_{11}, x_{21}°, and x_{31}° with y_{11}, and x_{12}°, x_{22}°, and x_{32}° with y_{12}. Among the six rows, at present, only the row x_{11} corresponds to an actual model, i.e. the Harrod-Domar model. At present, the remainder of the 5 rows represent those of potential models. Which potential models can actually be implemented? This must be resolved through reverse engineering later on. In particular, in order to realize the potential models with implication y_{12}, i.e. x_{12}°, x_{22}°, and x_{32}°, some form of technological progress must be incorporated as a new assumption.

4.3 Contexts and Utility Functions in Empirics Space

We first examine how the set of relevant observations $\mathcal{Z}$ is constructed. Obviously, it is context-dependent. What we mean by that is the relevance of an observation is dependent on the time and location of the researcher. What of importance is happening in society at that time and place? The environment plays an important role in the selection of observations, due to their relevance, and eventually in the selection of models though reverse engineering.

For example, the Empirics space in Figure 7 represents the economic events of significance for Solow in the 1950's. The $\mathcal{Z}$ axis gives observations that Solow might have been aware of. In the 1930's, in addition to z_1 (growing total GDP), observations z_2 to z_4 (significant unemployment, price instability, significant fluctuations) were of most relevance, as the great depression influenced the thinking of many economists. For instance, Harrod and Domar likely used the observation of the great depression to motivate their work. From the vantage point of the 1950's, observations z_2 to z_4 seem antiquated. At that time, in addition to z_1, observations z_5 to z_7 (stable employment, smooth factor price changes, smooth growth) were more relevant as descriptions of the economy.

To begin, at time $t-1 = 1955$, the Harrod-Domar model was prevalent in macroeconomics. This is illustrated in Figure 7. In Figure 8, we have altered the diagram in Empirics and Art space so that the domain of $f_{\mathbb{E}}$ is now subsets of $\mathcal{Y}$, called L_S^Y. Among the empirical predictions of the Harrod-Domar model is that GDP grows at a constant rate; see Branson (1972) chapter 18, particularly equations (18) and (20). In Figure 7, the Harrod-Domar model $\{a_1, a_2\}$ has a single fixed point corresponding to the black dot $\{\{a_1, a_2\}, \{z_1, z_2, z_3, z_4\}\}$ in *Art space*. In other words, the Harrod-Domar model supports the observation that the economy is growing in terms of total GDP. However, the Harrod-Domar model can support only this one among four observations that Robert Solow might have considered to be of importance in 1955. It is thus natural to assume that he was not satisfied with the state of economic growth theory in 1955. Hence, let us consider, in general, how to measure the degree of satisfaction or utility level of the state of the literature in question for research at a given time. In order to explain the utility function in detail, however, we first must discuss how to determine the relation in the Chu space $R_E = \{e_{ij}\}$.

Given a set of viable implications $Y_i^* \in L_S^Y$ and an observation $z_j \in \mathcal{Z}$, how do we determine when the observation z_j is consistent with the set of implications Y_i^*? To figure this out, we set up a probability model for *Empirics* *space*. We could set up a Bayesian model if we had information about the probability associated with the various subsets of implications or elements of L_S^Y, but we find it easier to use a classical approach to hypothesis testing at this point. Thus, we assume that for each $Y^* \in L_S^Y$, there is a probability measure π_{Y^*} defined on the (discrete) algebra generated by the set $\mathcal{Z}$.[25] That is, the probability of an observation given a set of implications is assumed to be known and exogenous.[26] We impose a restriction on π_{Y^*} in line with the consistency condition detailed in Section 3:

$$\begin{array}{l}\text{For all } Y_1^*, Y_2^*, Y_3^* \in L_S^Y \text{ and } z \in \mathcal{Z}, \text{ when } Y_3^* = Y_1^* \cap Y_2^* \\ \pi_{Y_3^*}(z) = \min\left\{\pi_{Y_1^*}(z), \pi_{Y_2^*}(z)\right\}\end{array}$$

We set a critical level $\overline{\pi}$, $0 \leq \overline{\pi} \leq 1$, and we define:

$$\begin{aligned} e_{ij} &= 1 \text{ when } \pi_{Y_i^*}(z_j) \geq \overline{\pi} \\ e_{ij} &= 0 \text{ when } \pi_{Y_i^*}(z_j) < \overline{\pi} \end{aligned}$$

Based on this definition, we set

$$Y_i^* R_E z_j \text{ if and only if } e_{ij} = 1$$

Next we introduce two utility functions. The first is a general utility function that applies to all of our diagrams, whereas the second is a utility function that applies to a particular state. Both take as the domain *Empirics space*. Let *Empirics space* be the Chu space $E = \{L_S^Y, R_E, \mathcal{Z}\}$, where $R_E = \{e_{ij}\}$. Then we define the *absolute* utility function as

$$U(E,t) = \sum_j w_j(t) \cdot \min\{\sum_i e_{ij}, 1\} \tag{7a}$$

$$\text{where } w_j(t) = \begin{cases} 1 \text{ for } z_j \in Z_{Solow}(t) \\ 0 \text{ for } z_j \in Z(t) - Z_{Solow}(t) \end{cases} \tag{7b}$$

and the *relative* utility that applies to a particular state as

$$u(E,t) = \frac{U(E,t)}{\#Z_{Solow}(t)} \tag{7c}$$

Here, $Z(t)$ represents the set of all observations at time t, whereas $Z_S(t)$ represents the subset of $Z(t)$ that is of relevance to Solow at time t. The notation $\#Z_{Solow}(t)$ represents the number of elements of set $Z_{Solow}(t)$. In the absolute

[25] In fact, since we never use joint probabilities, we could simply construct a separate algebra for each $z \in \mathcal{Z}$, including only it and its complement.

[26] A Bayesian approach would require a probability measure over $L_S^Y \times \mathcal{Z}$.

utility function, when $\sum_i e_{ij} \geq 1$, $\min\{\sum_i e_{ij}, 1\} = 1$, implying that observation j "has been explained by" or "is consistent with" *some* model. In this way, the *min-operation* is introduced to avoid double-counting in the evaluation of E. The absolute utility U is used to derive marginal utility for moving from 1955 to 1956, in other words, the dynamics. The relative utility u is used to calculate the satisfaction level of a researcher for a given state, say the one represented in Figure 7.[27]

Likewise, we define two (indirect) utility functions for each viable model $x_i \in \hat{\mathcal{X}}$. Setting $X = x_i$ in Section 3, let $Y_S^*(x_i) = \{y \in \mathcal{Y} \mid x_i R_S y\}$ be the set of implications associated with model $x_i \in \hat{\mathcal{X}}$. Then, we define the absolute utility and the relative utility of each viable model $x_i \in \hat{\mathcal{X}}$ as follows:

$$V(x_i, t) \equiv V(Y_S^*(x_i), t) = \sum_j w_j(t) \cdot e_{ij} \tag{7d}$$

$$v(x_i, t) \equiv v(Y_S^*(x_i), t) = \frac{V(x_i, t)}{\# Z_{Solow}(t)} \tag{7e}$$

Returning to Figure 7, $Z(t) = \{z_1, z_2, ..., z_7\}$ and

$$e_{1j} = \begin{cases} 1 \text{ for } j{=}1{,}2{,}3{,}4 \\ 0 \text{ otherwise} \end{cases}$$

whereas since $Z_{Solow}(t) = \{z_1, z_5, z_6, z_7\}$,

$$w_j(t) = \begin{cases} 1 \text{ for } j{=}1{,}5{,}6{,}7 \\ 0 \text{ otherwise} \end{cases}$$

and we have a single viable model, $x_1 = HD$, in Figure 7.

Hence, $U(E, t) = V(HD, t) = 1$ and $u(E, t) = v(HD, t) = \frac{1}{4}$. Given the latter, Solow must have been dissatisfied with the state of growth theory in 1955. Solow looked at the empirical implications of the Harrod-Domar model (in the *Empirics* quadrant), and found them in contradiction with reality at that time, for example the "growing unemployment or prolonged inflation" predicted by the Harrod-Domar model. So he wanted to construct a model that did not have this implication. Using reverse engineering, instead of the implications of the Harrod-Domar model, namely $\{y_1, y_2, y_3\}$, he wanted to construct a model that led to implications $\{y_1, y_4, y_5\}$ ={growing total GDP,

[27] There are many other ways of setting up the probabilities and utility functions. For example, we could use a Bayesian approach and/or an expected utility function, or we could use simultaneously multiple observations in $\mathcal{Z}$. The latter approach would require tracking contraries in observations, for instance. A Bayesian approach would mean that we couldn't apply the consistency condition. Here we are trying to keep the structure as simple as possible.

full employment, converging to balanced growth path}. To do so, he had to change at least one of the Harrod-Domar assumptions $\{a_1, a_2\}$. Since a_1 represents the usual assumptions of growth theory, the obvious candidate is a_2, namely replacing fixed factor proportions with something else. In the present context, the only natural replacement for fixed proportions is a_2^c = neoclassical production technology.

Now we move to Figure 8, and time $t = 1956$. This provides motivation for Solow to add $a_3 = a_2^c$ to the assumptions, and generates two new fixed points, namely $\{a_1, a_3\}$ = Solow 56 and $\{a_1\}$. That $\{a_1\}$ by itself is a fixed point means that any model incorporating the set of basic assumptions, a_1, can explain growing total GDP. Recall that a_1 includes the assumption that labor force increases at a constant relative rate. Thus, from the viewpoint of Solow, utility $U(E, t)$ rises from 1 to 4, which is equal to $U(\text{Solow 56 model}, t)$, the utility of the Solow 56 model. Furthermore, the relative utility rises from $\frac{1}{4}$ to 1, which again is equal to the relative utility of the Solow 56 model.

4.4 The Art of Reverse Engineering

To implement reverse engineering, we begin by reverse engineering from observations that we wish to account for back through Empirics space, then back though Science space, and finally back through Theory space. The composition of these inverse maps will achieve our goal.

In the implementation of reverse engineering, the utility function plays a very important role. The idea behind the use of utility functions is as follows. Although the utility functions are defined directly on Empirics space, what we search for is a model or a set of models that induces the highest utility score, or indirect utility, in Science space. Associated with each model in $\mathcal{X}$ is a set of implications in $\mathcal{Y}$, which in turn induce observations in $\mathcal{Z}$. We seek the model or set of models that induce the highest indirect utility score in Empirics space. Such a model might not be unique, and it might not induce all of the observations in $\mathcal{Z}$, just a strict subset.

For example, as we saw in Empirics space in Figure 7, the implications of the Harrod-Domar model have a low score from Solow's viewpoint. So Solow wanted to find a new model with a higher score. Reverse engineering back to Science space, please refer to Figure 12(b). By logically extrapolating Figure 12 (b), we can assign scores to each viable model in Science space. The model or models with the highest score are the ones Solow was interested in. In this case, potential model x_2° with the set of implications $\{y_1, y_2^c, y_3^c\}$ has the highest utility $U = 4$ and $u = 1$.

Reverse engineering from Science space to Theory space is much harder. How do we create a set of viable assumptions that generate the implications we desire? One way to proceed is to begin by listing all of the contraries of all the assumptions in $\mathcal{A}$. Assuming they are finite, in principle we may then list all possible combinations of these contraries with the list of original assumptions. Restrict attention to potentially viable models. Then we solve all such models, for example using a computer, choosing the one that yields highest utility. However, the number of such viable models can expand rapidly with the number of assumptions. Furthermore, this methodology requires a very precise, concrete specification of assumptions, including functional form assumptions. Therefore this approach may be inefficient and lack generality in practice. So we investigate an alternative methodology that may alleviate these difficulties.

To begin this alternative approach, we seek to determine what is possible and what is not possible in our framework. It is natural first to inquire about the limits of what we can do. Although the following theorem is rather obvious, it is still useful to state it formally:

Impossibility Theorem: Given $\mathcal{A}(t)$, suppose that the process of knowledge creation has reached a static knowledge state, and that the complemented lattice of fixed points is known. This implies that $f_{\mathbb{T}} : L_T^A \to L_T^X$, $f_{\mathbb{S}} : L_T^X \to L_S^Y$, and $f_{\mathbb{E}} : L_S^Y \to L_E^Z$ are known to the researcher. So the entire set of implications at this time is $f_{\mathbb{S}}(f_{\mathbb{T}}(L_T^A))$. Let $\mathcal{Y}^+$ be an extension of $\mathcal{Y}$ to a larger set of implications.

(i) Given $Y^* \in \mathcal{P}(\mathcal{Y}^+)$, if Y^* contains a pair of contraries, it is not viable and hence it cannot be explained by a single model or set of assumptions.

(ii) Given $Y^* \in \mathcal{P}(\mathcal{Y}^+)$, if $Y^* \notin f_{\mathbb{S}}(f_{\mathbb{T}}(L_T^A))$, then Y^* cannot be explained by any viable model based on the current set of assumptions. Thus, the researcher must look for a new set of assumptions and a new model to explain Y^*.

For example, in Figure 7, the Harrod-Domar model represents a static knowledge state that has implications $Y = \{y_1, y_2, y_3\}$, whereas $L_T^A = \{\{a_1, a_2\}\}$. Let $Y^* = \{y_1, y_4, y_5\}$ be another subset of the extended set of implications, for instance those in Figure 8. Since $Y^* \notin f_{\mathbb{S}}(f_{\mathbb{T}}(L_T^A))$, any model based on the existing set of assumptions $\{a_1, a_2\}$ cannot explain Y^*.

A more concrete example can be found in Figures 8 and 9. The Harrod-Domar model and the Solow (1956) model exhaust all viable models based on assumptions $\{a_1, a_2, a_3\}$, where $a_3 = a_2^c$. Therefore, to obtain any implication that is not a specification of or not included in the implications $\{y_1, y_2, ..., y_5\}$, more assumptions and a new model must be used. For example, the implication "growing per capita income" requires new assumptions.

Next we search for possibility results. These are not necessarily as general as the impossibility theorem, but possibility theorems will, at a minimum, give us a direction for further constructive work. To head in this direction, we must specialize our framework. This will involve setting up a Euclidean space of parameters to be used in conjunction with assumptions, and a Euclidean space of implications. To proceed, we isolate a set of assumptions upon which we focus. We impose more structure on these assumptions, conditional on the other assumptions we make for a particular model. Suppose that we have a parameter space, P, for these isolated assumptions. We assume that P is contained in the closure of an open subset of a Euclidean space. In parallel, we suppose that we have a space of implications $\mathbb{R}^l$.

Returning to our context of early growth theory, we begin with an example. The dynamics of the early models can be expressed by a single equation:

$$\dot{K} = sF(K, L)$$

where s is the exogenous saving ratio. Further assume that labor supply increases exponentially with time:

$$\frac{\dot{L}}{L} = n$$

where n is the exogenous rate of increase in labor supply. Suppose that F satisfies constant returns to scale. Then if we define the per capita variable $k = K/L$ and $f(k) = F(K, L)/L = F(k, 1)$, the dynamics can be rewritten as follows:

$$\dot{k} = sf(k) - nk$$

Let r represent the rental rate of capital and let ω be the wage. Retaining the assumption of competitive factor markets and normalizing the price of output to be 1, wherever the derivative $f'(k)$ exists,

$$\begin{aligned} r &= f'(k) \\ \omega &= f(k) - k \cdot f'(k) \end{aligned}$$

In this context, for the Harrod-Domar model, taking units of K and L appropriately,

$$F(K, L) = \min\{K, L\}$$

Then the dynamics become

$$\dot{k} = s \min\{k, 1\} - nk$$

To place the Harrod-Domar model in a more general context, let us parameterize the production functions available for the early growth models using the CES function:

$$F(K, L) = \left[\alpha K^\beta + (1-\alpha)L^\beta\right]^{1/\beta}, \quad 0 < \alpha < 1, \quad -\infty < \beta < 1 \tag{8}$$

In this case, the model is parameterized by the production function, so $P = (0, 1) \times (-\infty, 1)$. The Harrod-Domar model corresponds to[28]

$$\lim_{\beta \to -\infty} \left[\alpha K^\beta + (1-\alpha)L^\beta\right]^{1/\beta} = \min\{K, L\} \tag{9}$$

The Cobb-Douglas case, namely the Swan (1956) model, corresponds to

$$\lim_{\beta \to 0} \left[\alpha K^\beta + (1-\alpha)L^\beta\right]^{1/\beta} = K^\alpha \cdot L^{1-\alpha} \tag{10}$$

In terms of the per capital variable $k \equiv K/L$, each equation above can be rewritten as follows:

$$f(k) = \left[\alpha k^\beta + (1-\alpha)\right]^{1/\beta}, \; 0 < \alpha < 1, -\infty < \beta < 1 \tag{11}$$

$$\lim_{\beta \to -\infty} \left[\alpha k^\beta + (1-\alpha)\right]^{1/\beta} = \min\{k, 1\} \tag{12}$$

$$\lim_{\beta \to 0} \left[\alpha k^\beta + (1-\alpha)\right]^{1/\beta} = k^\alpha \tag{13}$$

For later use, we calculate factor prices for the CES specification. For $-\infty < \beta < 1$,

$$r = f'(k) = \alpha \cdot \left[\alpha k^\beta + (1-\alpha)\right]^{(1-\beta)/\beta} \cdot k^{\beta-1} > 0 \tag{14}$$

$$w = f(k) - kf'(k) = (1-\alpha) \cdot \left\{\alpha k^\beta + (1-\alpha)\right\}^{(1-\beta)/\beta} \cdot k^{\beta-1} > 0 \tag{15}$$

and hence

$$\frac{w}{r} = \frac{1-\alpha}{\alpha} \cdot k^{1-\beta}, \quad -\infty < \beta < 1 \tag{16}$$

In the case of the Harrod-Domar specification (9) or (12), factor prices are

$$r = \begin{cases} 1 \text{ for } k<1 \\ 0 \text{ for } k>1 \end{cases} \qquad \text{(when } \beta = -\infty) \tag{17}$$

$$w = \begin{cases} 0 \text{ for } k<1 \\ 1 \text{ for } k>1 \end{cases} \qquad \text{(when } \beta = -\infty) \tag{18}$$

[28] We could parameterize the models in another way, for example using $\frac{1}{\beta}$ in place of β. But in either case, infinity must appear somewhere for the representation of the Harrod-Domar model, so we choose a convenient representation. For the proof of (9) and (10), see Varian (1978).

but the factor prices are undefined for $k = 1$. This discontinuity of factor prices derives from the kink in the isoquants.

Solow looked at the state of the literature and wanted to achieve a qualitative change in implications, given his absolute and relative utility functions. Specifically, although the Harrod-Domar model implied expanding unemployment of one factor, Solow wanted a model that always features full employment of both factors. In correspondence with this goal, likely a radical change would be needed to obtain a qualitative change in implications. Thus, movement away from $-\infty$, essentially by adding a new dimension, β, to the parameter space, was needed. As it turned out, this additional dimension could accomplish the goal. This led to Solow (1956).

Next we discuss singularities, the mathematical interpretation of the preceding discussion of economics. In preparation for the sequel, it is useful to recall a standard mathematical definition. A real analytic function h with domain and range subsets of Euclidean spaces has an *essential singularity at point* $\overline{s}$ in its domain if and only if

$$\lim_{s \to \overline{s}} h(s) \text{ and } \lim_{s \to \overline{s}} \frac{1}{h(s)} \text{ do not exist.}$$

Consider the Harrod-Domar model, where $\beta = -\infty$. We consider f_S to be a composition of maps, the first of which is given in equation (8). This first map is well-behaved at $\beta = -\infty$. But eventually f_S will map from the parameter space through the production function to equilibrium prices, given for example by the equilibrium factor price ratio (16). From here, equilibrium implications, such as price stability or instability, will be derived. Choose parameters so that equilibrium along the balanced growth path has $k \neq 1$. Notice that the standard definition of an *essential singularity at infinity* for a map $f_S(b)$ is an essential singularity of the map $f_S(1/b)$ at $b = 0$. Substituting this into (16), we can see immediately that since $k^{1-\frac{1}{b}}$ has an essential singularity at $b = 0$

for any fixed $k \neq 1$,[29] $f_S(\beta)$ has an essential singularity at $\pm\infty$.[30] Thus, the Harrod-Domar model, represented by $\beta = -\infty$, is an essential singularity in the parameter space.

This is the mathematical essence of what Solow (1956) accomplished: It moved away from an essential singularity in the parameter space, where even a small move would yield radically different results.

Helpful Hint 1: To achieve a qualitative change in implications, either make a qualitative change in assumptions or find a larger parameter space in which the current assumptions represent an essential singularity, and move away from it.

In other words, suppose that we have a map from the parameter space of assumptions to a (Euclidean) space of implications. Then, to achieve a qualitative change in implications, we need either a radical change in assumptions or a move away from an essential singularity in the map.

A second, less radical example is the move from Solow (1956) to Solow (1957). Here, Solow (1956) had a model with constant per capita GDP growth on the balanced growth path, but wanted a model with an implication of growing per capita GDP with all the other implications maintained. So Solow (1957) introduced a new shift parameter $a \geq 0$ into the production function:

$$e^{a\tau} \cdot F(K, L) \tag{19}$$

[29] Actually, since we are performing comparative statics, the proper expression here is $\frac{w}{r} = \frac{1-\alpha}{\alpha}(k^*)^{1-\beta}$, where k^* is the equilibrium capital-labor ratio. In the CES case, we can compute k^* as

$$k^*(\alpha, \beta) = \frac{(1-\alpha)^{\frac{1}{\beta}} \cdot s}{(n^\beta - \alpha \cdot s^\beta)^{\frac{1}{\beta}}}$$

Hence,

$$\begin{aligned} \frac{w}{r} &= \frac{1-\alpha}{\alpha} \cdot k^*(\alpha, \beta)^{1-\beta} \\ &= \frac{(1-\alpha)^{\frac{1}{\beta}}}{\alpha} \cdot \frac{1}{[(\frac{n}{s})^\beta - \alpha]^{\frac{1}{\beta}-1}} \end{aligned}$$

To check for an essential singularity at $\beta = \pm\infty$, we substitute $\beta = \frac{1}{b}$ and let $b \to 0$ (or alternatively we can reparameterize the system by b) and look at the limits as we approach 0 from above and below. The limit is the same as the limit of $\frac{1}{\alpha}[(\frac{n}{s})^{\frac{1}{b}} - \alpha]$. Therefore, as long as the exogenous parameters $n \neq s$, we have an essential singularity at $\beta = \pm\infty$ (or $b = 0$). In other words, the limit of $\frac{1}{\alpha}[(\frac{n}{s})^{\frac{1}{b}} - \alpha]$ is different as $b \to 0$ from above and below, and the same is true of its inverse. In fact, we could have simply focused on k^* itself to see the essential singularity.

[30] In fact, the classical example of an essential singularity at zero is $\exp(\frac{1}{b})$.

where τ represents time in the model. Choosing $a > 0$ naturally leads to per capita GDP growth on the balanced growth path. The implications of the models are continuous in the parameter a so that Solow (1956) is essentially a special case of Solow (1957). In other words, a small change in implications does not require a qualitative change in assumptions.

In order to state the next theorem, we need some preparation. Analogous to P, let $\mathbb{R}^l$ (where l is not necessarily the same as the dimension of P) represent the implications of a model upon which the researcher is focused. The initial set of models or assumptions is represented by a set $P' \subseteq P$. For example, it could be the set of parameters where $a = 0$ in the last example. The set of implications could include all implications that feature constant per capita growth in GDP along the balanced growth path. So instead of the map $f_{\mathbb{S}} : L_T^X \to L_S^Y$, we consider a special case where $f_{\mathbb{S}} : P' \to \mathbb{R}^l$. Then we wish to address the following question: When can the Science map f_S from sets of assumptions or models to implications be extended from the current set of parameters P' to all of P? And when is this extension unique? This is one way to make progress in science. In mathematical terms, it reduces to: When can the science map can be extended uniquely?[31]

Theorem 2:[32] *Where the map from the initial assumptions to implications is sufficiently smooth, then for any additional parameters, there is a unique smooth extension of the map to the larger new parameter space, so the implications of a new model can be derived as an extension of the implications of the old models.*

For example, this theorem would apply to the move from Solow (1956) to Solow (1957). There, the new parameter a is added to the parameter space, and the new implication, growing per capita GDP, follows naturally. This might seem trivial when the new implication is easy to find. But there are also cases when the analytic extension is not so easy to compute or obvious. Then it is useful to know that the new implications will be smoothly related to the new parameters and be unique.

More generally, as discussed in section 4.2 above, y_{11} (growing total GDP but not per capita) and y_{12} (growing total and per capita GDP) are *specifica-*

[31] Notice that if instead of extending P' to P, we instead extend it to $\overline{P}$ (for example through a change of variables from P to $\overline{P}$), the analytic extension will not be the same function on both of these domains.

[32] A formal statement and proof of Theorem 2 can be found in the Appendix.

tions or a *partition* of the existing implication y_1 (growing total GDP). We can expect that such specification or a partition of an existing implication can be achieved by introducing a smooth extension of an existing assumption or function, without involving a singularity. In contrast, when we move from one set of viable implications, say Y, to another set of viable implications Y^* containing at least one contrary to an implication in Y, then a singularity is involved. That is how we relate our discrete mathematics of logic to the smooth mathematics of analytic functions.

Notice that Helpful Hint 1 is essentially the contrapositive of Theorem 2.

We can relate our larger structure back to Kuhn (1962). An example of "normal science" is the move from Solow (1956) to Solow (1957), where the mathematics behind the innovation (but not the research itself) is the unique extension of a function, as given in Theorem 2. Though we know it exists and is unique, calculation of the extension can still be difficult. In contrast, the move from the Harrod-Domar model to Solow (1956) represents a "paradigm shift." The mathematics behind our description of the innovation (but not the research itself) is the introduction of a new parameter, β, into the parameter space so that the researcher can exit from a singularity point that cannot be removed. Once outside this singularity, the implications of models look very different from those generated using the singular model. Unlike the situations with no essential singularity, *the seeds of the extension are not present in the map f_S before the extension.*

A further example of a "paradigm shift" is the move from the framework of perfect competition to the framework of monopolistic competition with product differentiation and consumer preference for variety. In this example, in place of homogenous consumption goods, composite consumption is represented by C and its components consisting of differentiated consumption goods are represented by c_i as follows:

$$C = \{c_1^\rho + c_2^\rho + ... + c_n^\rho\}^{\frac{1}{\rho}} \tag{20}$$

where $0 < \rho < 1$. The elasticity of substitution between commodities is given by $\sigma = \frac{1}{1-\rho}$. The case where $\rho = 1$ is the case of no product differentiation, namely where the consumption goods are perfectly substitutable, corresponding to homogeneous products. This case reflects perfect competition with $\sigma = \infty$. So moving away from $\rho = 1$ requires moving away from perfect competition, suggesting another big "paradigm shift." To see this in more detail, in these models there is generally a positive fixed cost of production for each firm; each firm produces one particular variety of the commodity. At

$\rho = 1$, namely under perfect substitutes and perfect competition, no firm can cover its fixed cost at equilibrium. Thus, the only competitive equilibrium has no production. The map from ρ to equilibrium as $\rho \to 1$ from below is discontinuous. It is hard to discuss singularities more formally in this context, since the model with $\rho > 1$ or $\sigma < 0$ involves nonconvexities and product differentiation.

This analysis should offer a caution to researchers using models that rely heavily on specific functional forms in conjunction with parameters. Do the models represent an essential singularity in a larger parameter space? In other words, are the implications robust?

Beyond the study of extensions of functions, the framework we have developed has further properties. Chief among these is an "all or (almost) nothing" property, detailed next:

Theorem 3:[33] *Where the map from the parameters to implications is sufficiently smooth, consider any dimension of the implication space* $\mathbb{R}^l$*, fixing the other coordinates. Then either the same implication holds for all values of the parameters mapping into the restricted implications set, or the set of parameters having any particular implication in this set is of measure zero.*

A nice example of this is Solow (1957). What Theorem 3 says is that either per capita GDP is growing at the same rate for all parameters, or no given level of per capita GDP growth holds for a positive measure of parameters.

5 The Dynamic Process of Knowledge Creation

In this section, by synthesizing the aspects of our framework that we have detailed, we examine the dynamic process of knowledge creation for the one agent case.

To begin, there is a mathematical structure providing underpinnings for our examples. Notice that we are using a combination of discrete mathematics, for example Boolean algebras, and very smooth mathematics, namely analytic functions. This seems essential, as do the connections between these two types of mathematics. To quote von Neumann (1958, p. 75):

[33] A formal statement and proof of Theorem 3 can be found in the Appendix.

> Any artificial automaton that has been constructed for human use, and specifically for the control of complicated processes, normally possesses a purely logical part and an arithmetical part, i.e., a part in which arithmetical processes play no role, and one in which they are of importance. This is due to the fact that it is, with our habits of thought and of expressing thought, very difficult to express any truly complicated situation without having recourse to formulae and numbers.

So how do we form the formal connection? In our framework, it is the singularities in the analytic functions, essential or not, that allow discrete jumps in Science space between implications, namely between elements of $\mathcal{Y}$. In an example of the previous section, there is an essential singularity at $\beta = \pm\infty$ for all $\alpha \in (0,1)$. Thus, the entire parameter space $P' = (0,1) \times \{-\infty\}$ is essentially singular in $P = (0,1) \times [-\infty, \infty]$. This is how we find a discrete move in the lattice L_S^Y from a viable implication set Y to a new viable implication set Y^* containing one of its contraries (as opposed to a specification) in $\mathcal{Y}^+$. If Y^* does not contain a contrary of Y, then the move to the new implications set is smooth, and can be found through the (unique) analytic extension.

(To Be Continued)

6 Extension

Having explored how work with positive questions and issues is accomplished in our diagrams, we next turn to the analysis of normative questions and issues. For this analysis, we refer to Figure 16. We reproduce our diagram for positive analysis in the upper portion of Figure 16, complete with all of our working spaces. In the lower part of Figure 16, we place our new working space, the one corresponding to normative working space. We describe each axis in turn.

First we describe the elements of the axis $\widetilde{\mathcal{A}}$. For normative analysis, this will consist of the objectives and constraints that are available to the researcher. Next, the elements of the axis $\widetilde{\mathcal{X}}$ are the permissible models, each of which consists of an objective function and the associated set of permissible constraints. The axis $\widetilde{\mathcal{Y}}$ contains the possible normative implications of permissible models. Finally, the axis $\widetilde{\mathcal{Z}}$ contains as elements the policy implications of the analysis of the models. For example, $\widetilde{\mathcal{Z}}$ might contain various taxes, such as Pigouvian taxes (in the case of externalities) or Ramsey taxes (in the case of taxing private goods), or various types of regulation.

The maps in normative theory space and in normative science space are analogous to the corresponding maps in positive working space.

7 Appendix: Formal Statements and Proofs of Theorems 2 and 3

To be more precise mathematically, we must introduce some definitions before stating the main theorem on dynamics. A set D' is called *real analytic* if it is a closed subset of a Euclidean space and each of its points has a neighborhood U such that $U \cap D'$ is the set of common zeroes of a finite number of analytic functions in U.

Theorem 2: Let P' be a closed real analytic subset of a connected, open subset P of a Euclidean space. Let $f_{\mathbb{S}} : P' \to \mathbb{R}^l$ be analytic. If there exist analytic functions $g_1, ..., g_k$ on P such that $P' = \{p \in P \mid g_1(p) = g_2(p) = ... = g_k(p) = 0\}$, then there is an analytic extension of f_S to P, namely $\widehat{f}_S : P \to \mathbb{R}^l$. Moreover, if P' contains a converging sequence of distinct points (and consequently their accumulation point), then the analytic extension to P is unique.

Proof of Theorem 2: Existence of an extension follows from Nardelli and Tancredi (1996), Proposition 1 and the discussion in the introduction. Uniqueness follows from the Identity Theorem of complex analysis.

Of course, a special case of the major assumption of the theorem is where $k = 1$ and $g_1(p)$ is simply the projection of p onto one of its components. In this case, P' is called *principal.* In contrast with existence and uniqueness of an analytic extension, actually computing one takes some work.

Theorem 3: Let P be an open and connected subset of a Euclidean space, and let $\widehat{f}_S : P \to \mathbb{R}$, where $\widehat{f}_S$ is analytic. Then either $\widehat{f}_S$ is constant, or the Lebesgue measure of the set in P that has any given value in $\mathbb{R}$ is zero.

Proof of Theorem 3: Follows directly from Rader (1979), Lemma 4.

References

[1] Barwise, J. and J. Seligman, 1997. *Information Flow.* Cambridge: Cambridge University Press.

[2] Berge, C., 1963. *Topological Spaces.* New York: Macmillan Company.

[3] Branson, W.H., 1972. *Macroeconomic Theory and Policy.* New York: Harper and Row.

[4] Burmeister, E. and A.R. Dobell, 1970, *Mathematical Theories of Economic Growth*, London: The Macmillan Company.

[5] Dimand, R.W. and B.J. Spencer, 2008. "Trevor Swan and the Neoclassical Growth Model." NBER Working Paper 13950.

[6] Domar, E., 1946. "Capital Expansion, Rate of Growth, and Employment." *Econometrica* 14 (2): 137–147.

[7] Dunn, J.M. and G.M. Hardegree, 2001, *Algebraic Methods in Philosophical Logic*, Oxford: Oxford University Press.

[8] Ganter, B. and R. Wille, 1999. *Formal Concept Analysis: Mathematical Foundations.* Berlin: Springer.

[9] Ganter, B., G. Stumme, and R. Wille (eds.), 1998. *Formal Concept Analysis.* Lecture Notes in Artificial Intelligence 3626, Subseries of Lecture Notes in Computer Science. Berlin: Springer.

[10] Harrod, R.F., 1939. "An Essay in Dynamic Theory." *The Economic Journal* 49 (193): 14–33.

[11] King, R.D., K.E. Whelam, F.M. Jones, P.G.K. Reiser, C.H. Bryant, S.H. Muggleton, D.B. Kell and S.G. Oliver, 2004, "Functional genomic hypothesis generation and experimentation by a robot scientist," *Nature* 427, 247-252.

[12] Kuhn, T.S., 1962. *The Structure of Scientific Revolutions*, Chicago: The University of Chicago Press.

[13] Kurzweil, R., 2005, *The Singularity Is Near*, New York: PENGUIN BOOKS.

[14] Nardelli, G. and A. Tancredi, 1996. “A Note on the Extension of Analytic Functions off Real Analytic Subsets.” *Revista Mathemática de la Universidad Complutense de Madrid* 9 (1): 85-97.

[15] Pratt, V., 1999. *Chu Spaces.* Notes for the School on Category Theory and Applications, University of Coimbra. http://chu.stanford.edu/guide.html#coimbra

[16] Rader, T., 1979. “Nice Demand Functions - II.” *Journal of Mathematical Economics* 6: 253-262.

[17] Rott, H., 2001. *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning.* Oxford Logic Guides 42. Oxford: Oxford University Press.

[18] Soldatova, L.N., A. Clare, A. Sparkes and R.D. King, 2006, “An ontology for a robot scientist,” *Bioinformatics* 22, e464-e471.

[19] Solow, R.M., 1956. “A Contribution to the Theory of Economic Growth.” *Quarterly Journal of Economics* 70, 65-94.

[20] Solow, R.M., 1957. “Technical Change and the Aggregate Production Function.” *Review of Economics and Statistics* 39, 312-320.

[21] Solow, R.M. and W.S. Vickrey, 1971. “Land Use in a Long Narrow City.” *Journal of Economic Theory* 3, 430-447.

[22] Stiglitz, J.E. and H. Uzawa (eds.), 1969. *Readings in the Modern Theory of Economic Growth*, Cambridge: The MIT Press.

[23] Swan, T., 1956. “Economic Growth and Capital Accumulation.” *Economic Record* 32, 334-361.

[24] Turing, A.M., 1936, “On Computable Numbers, with an Application to the Entscheidungsproblem.” *Proceedings of the London Mathematical society* 42: 230-65.

[25] Varian, H.R., 1978, *Microeconomic Analysis*, New York: W.W. Norton & Company.

[26] Villaverde, A.F. and J.R. Banga, 2014. “Reverse Engineering and Identification in Systems Biology: Strategies, Perspectives and Challenges.” *Interface* 11: 20130505.http://dx.doi.org/10.1098/rsif.2013.050

[27] von Neumann, John, 1958 (the 3rd edition in 2012), *The Computer & the Brain*, New Haven: Yale University Press.

[28] Watanabe, S., 1969. *Knowing and Guessing: A Qualitative Study of Inference and Information.* New York: John Wiley and Sons.

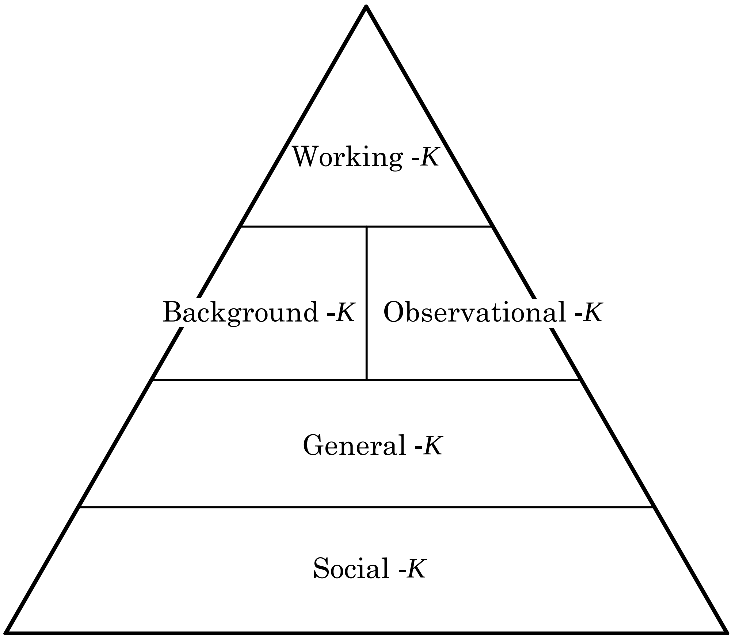


Figure 1: The Pyramid of Knowledge

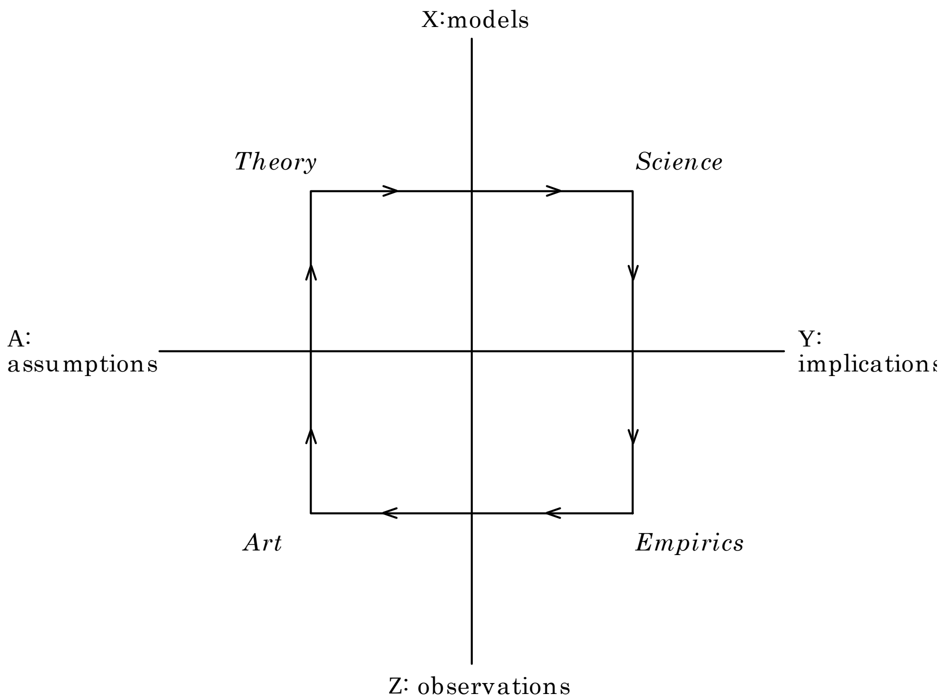


Figure 2: Working-K Space

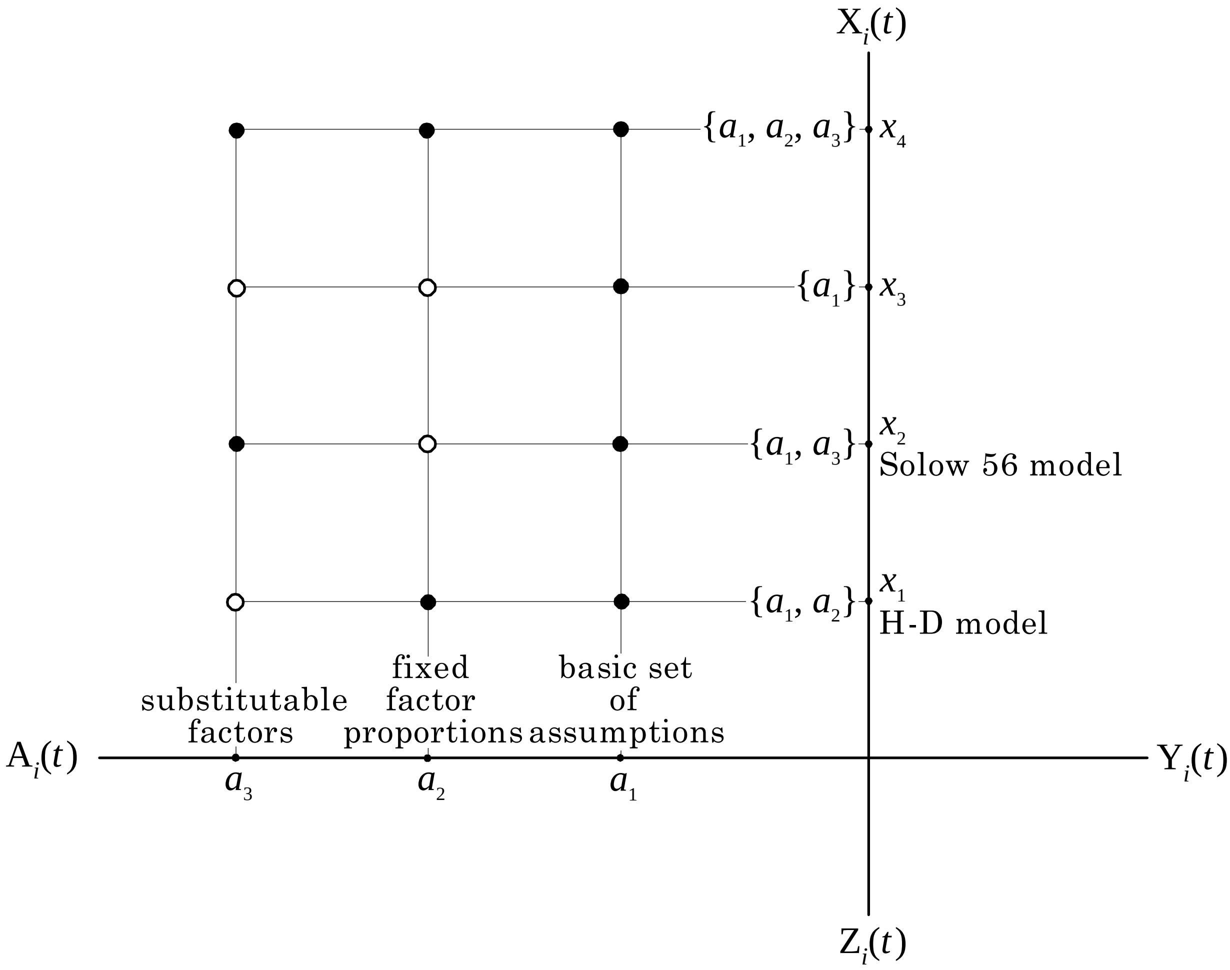


Figure 3: How to Work in *Theory Space*

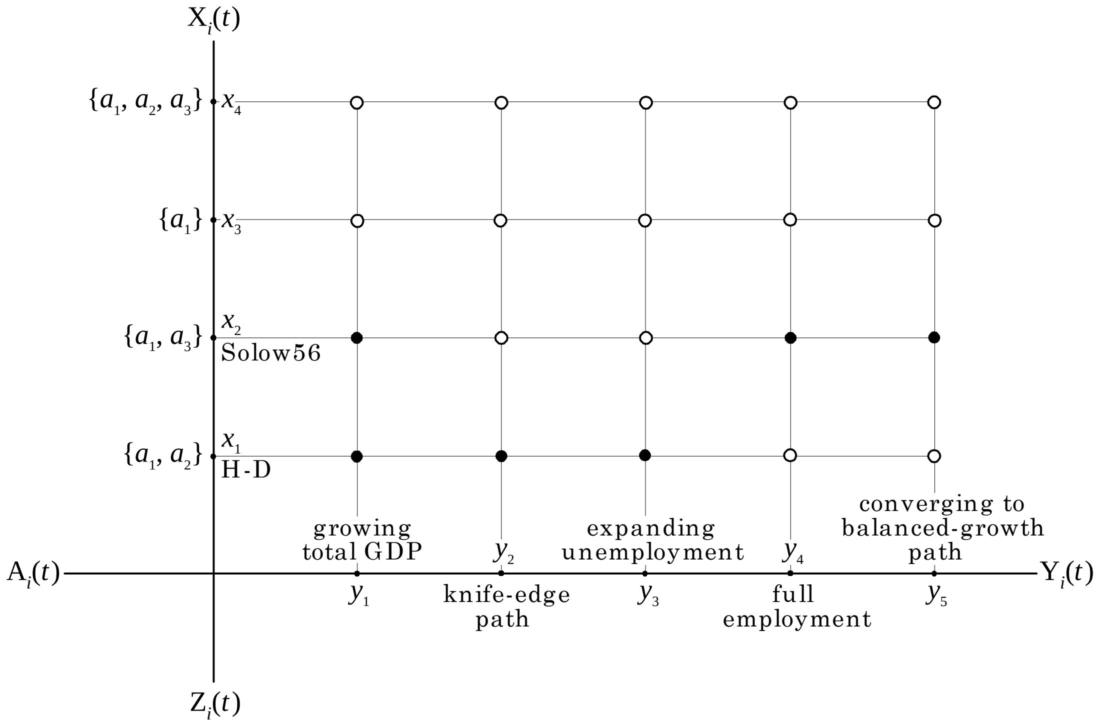


Figure 4: How to Work in *Science Space*

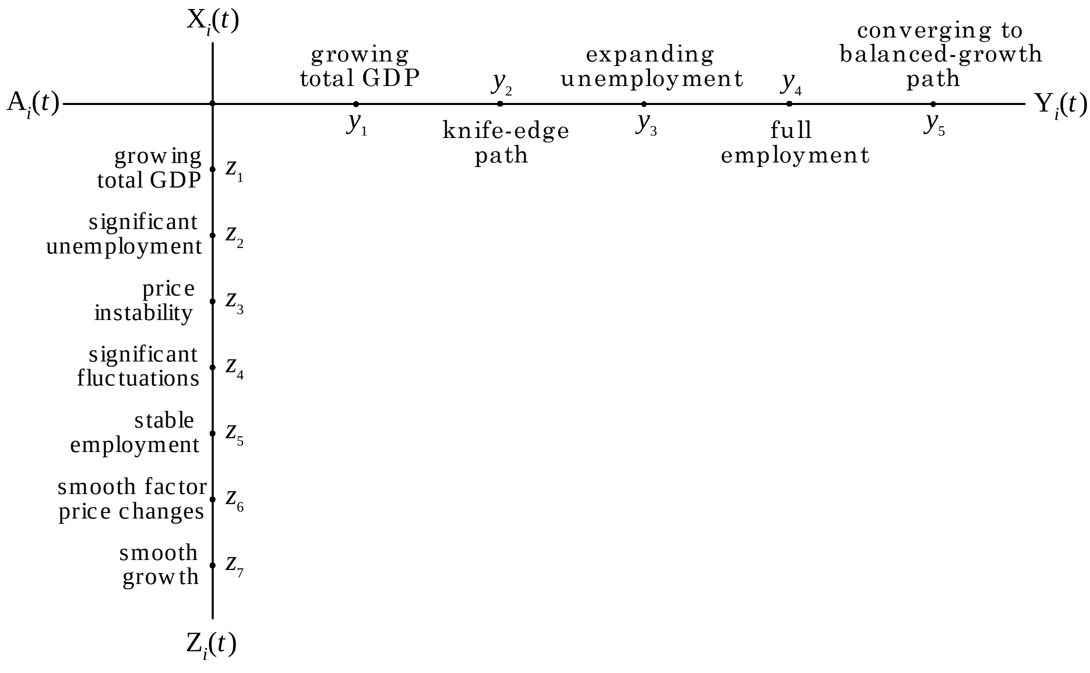


Figure 5: How to Work in *Empirics Space*

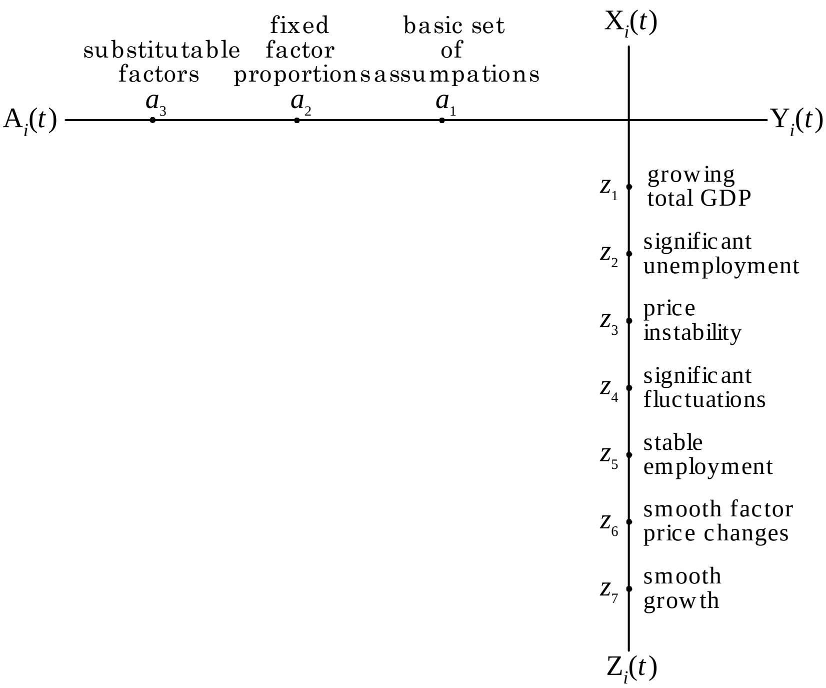


Figure 6: How to Work in *Art Space*

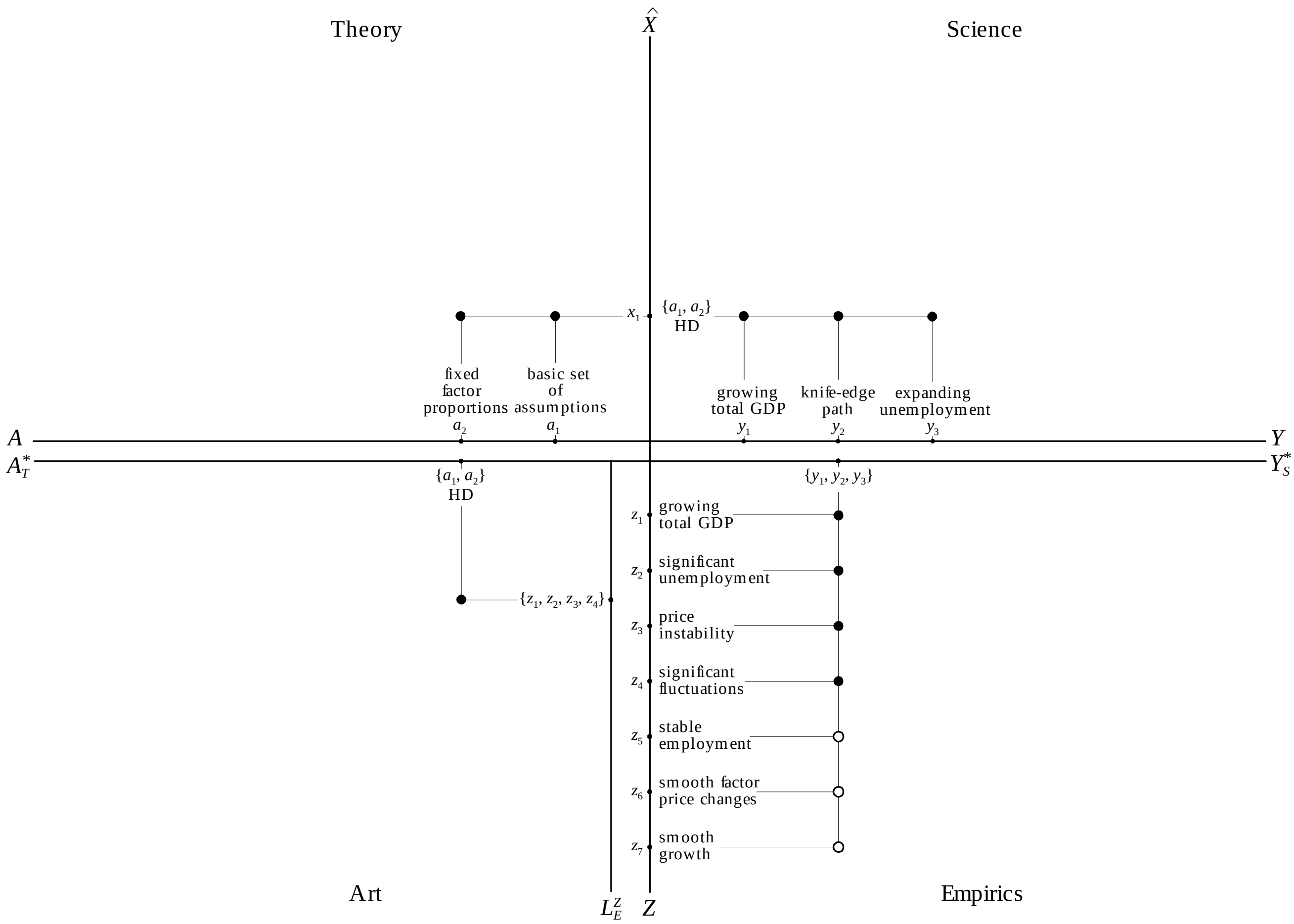


Figure 7: The Evolution of Growth Theory à la Robert Solow - Phase 1 (1955)

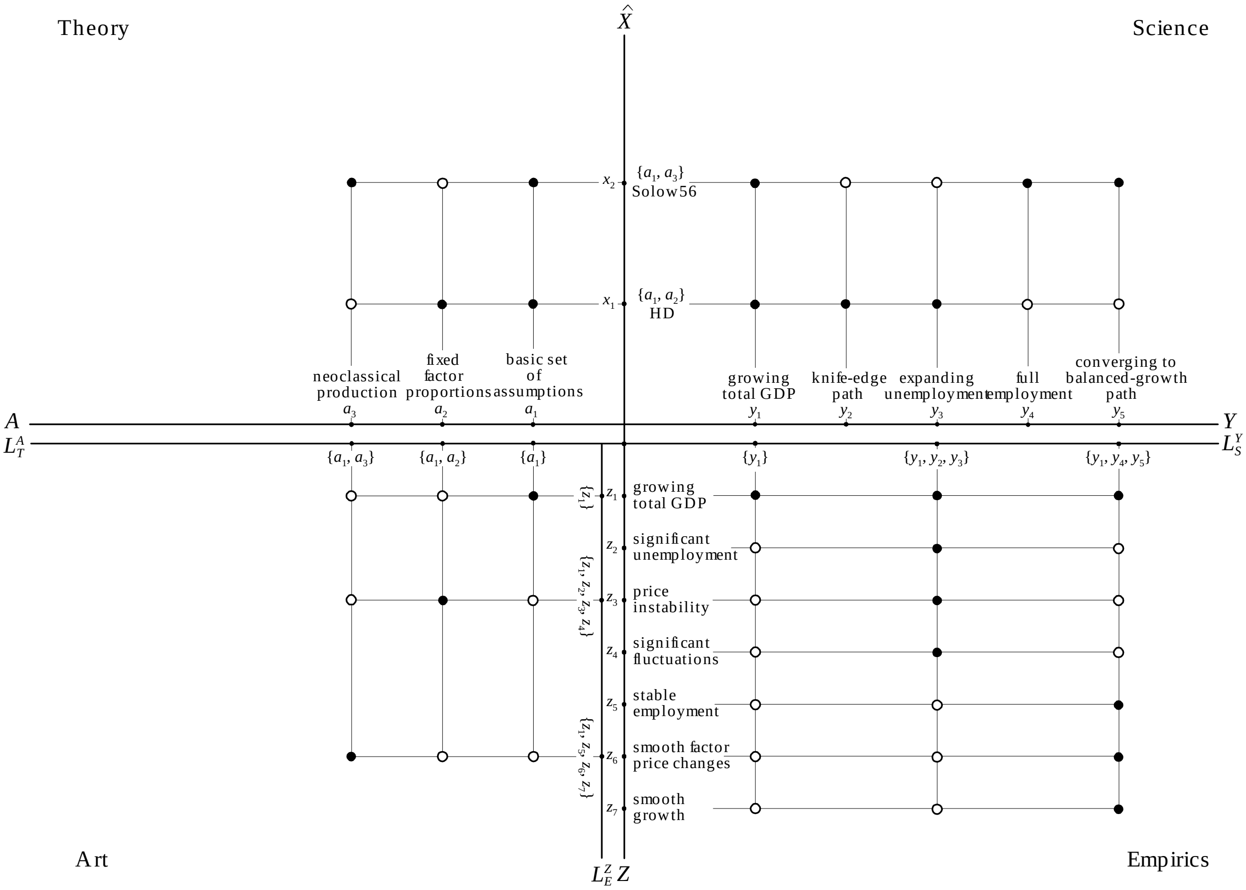


Figure 8: The Evolution of Growth Theory à la Robert Solow - Phase 2 (1956)

(a) Boolean algebras (b) Viable models and formal concepts

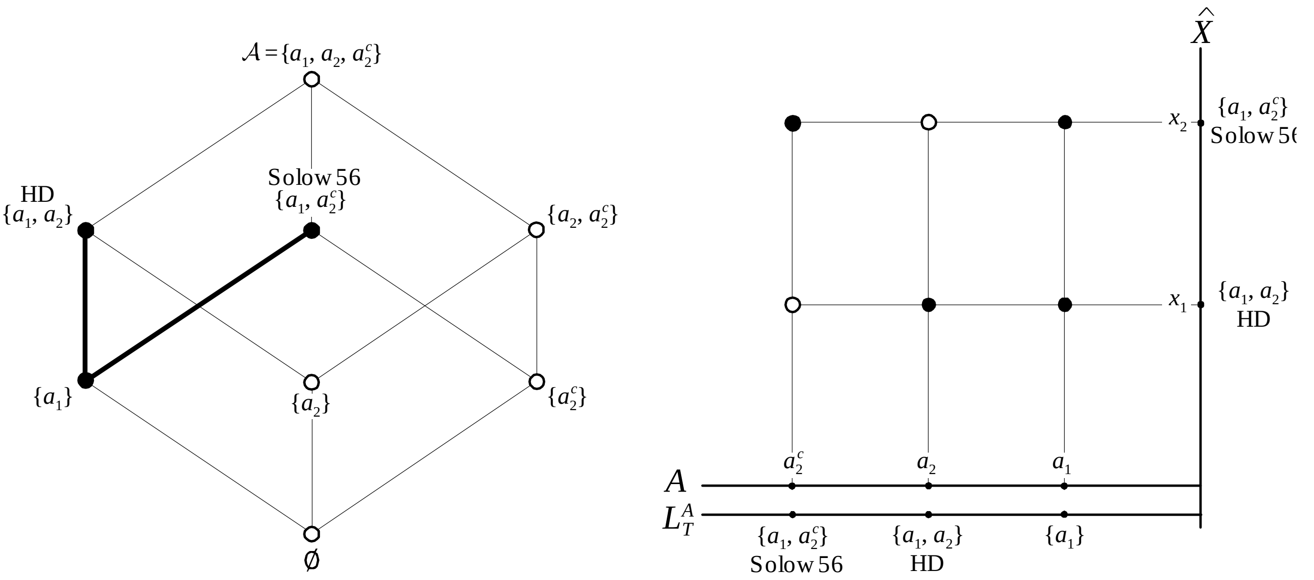


Figure 9: Boolean Algebras and Formal Concepts in Theory Space with $\mathcal{A} = \{a_1, a_2, a_2^c\}$

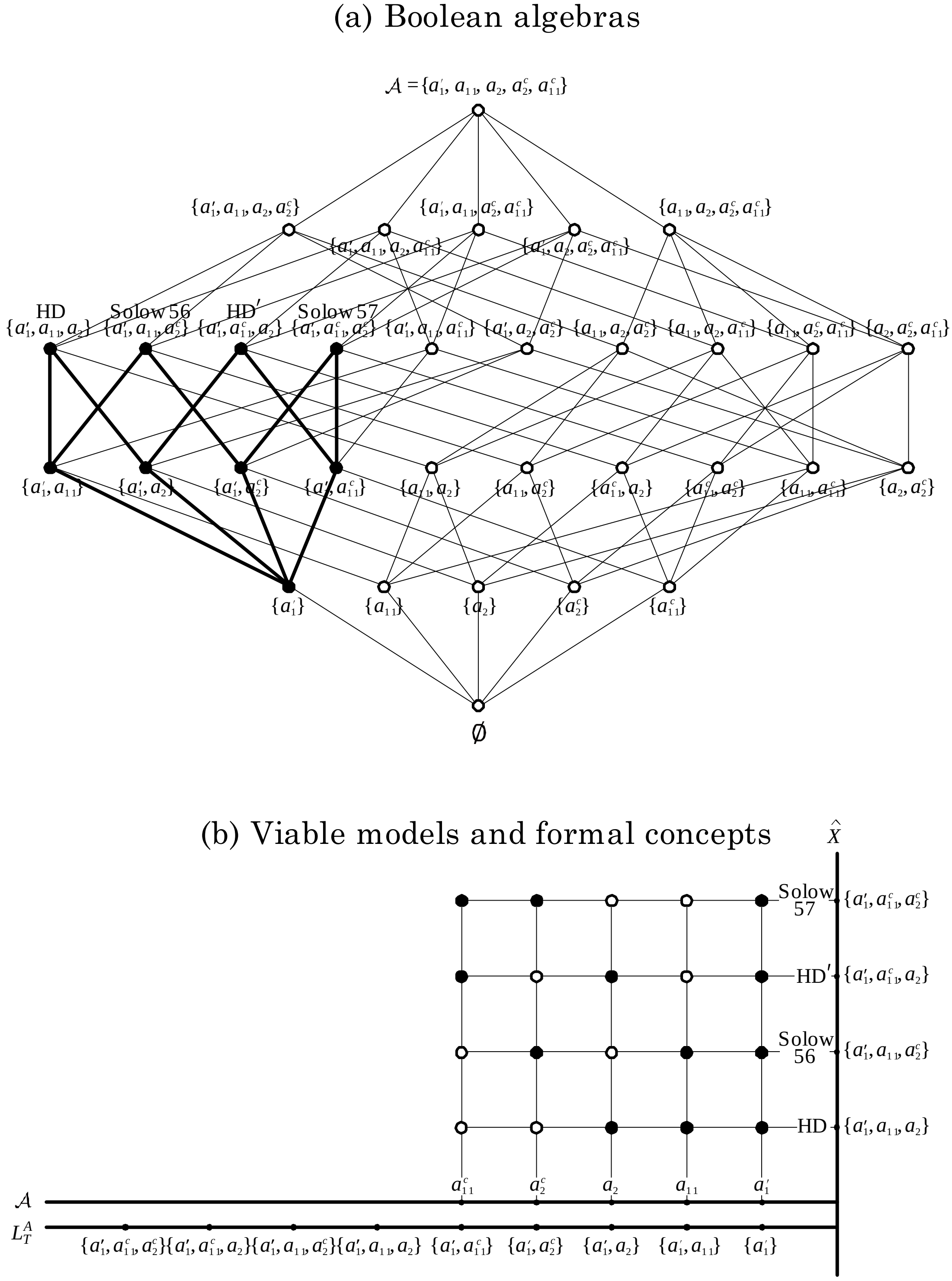


Figure 10: Boolean Algebras and Formal Concepts in Theory Space with $\mathcal{A} = \{a'_1, a_{11}, a_2, a^c_2, a^c_{11}\}$

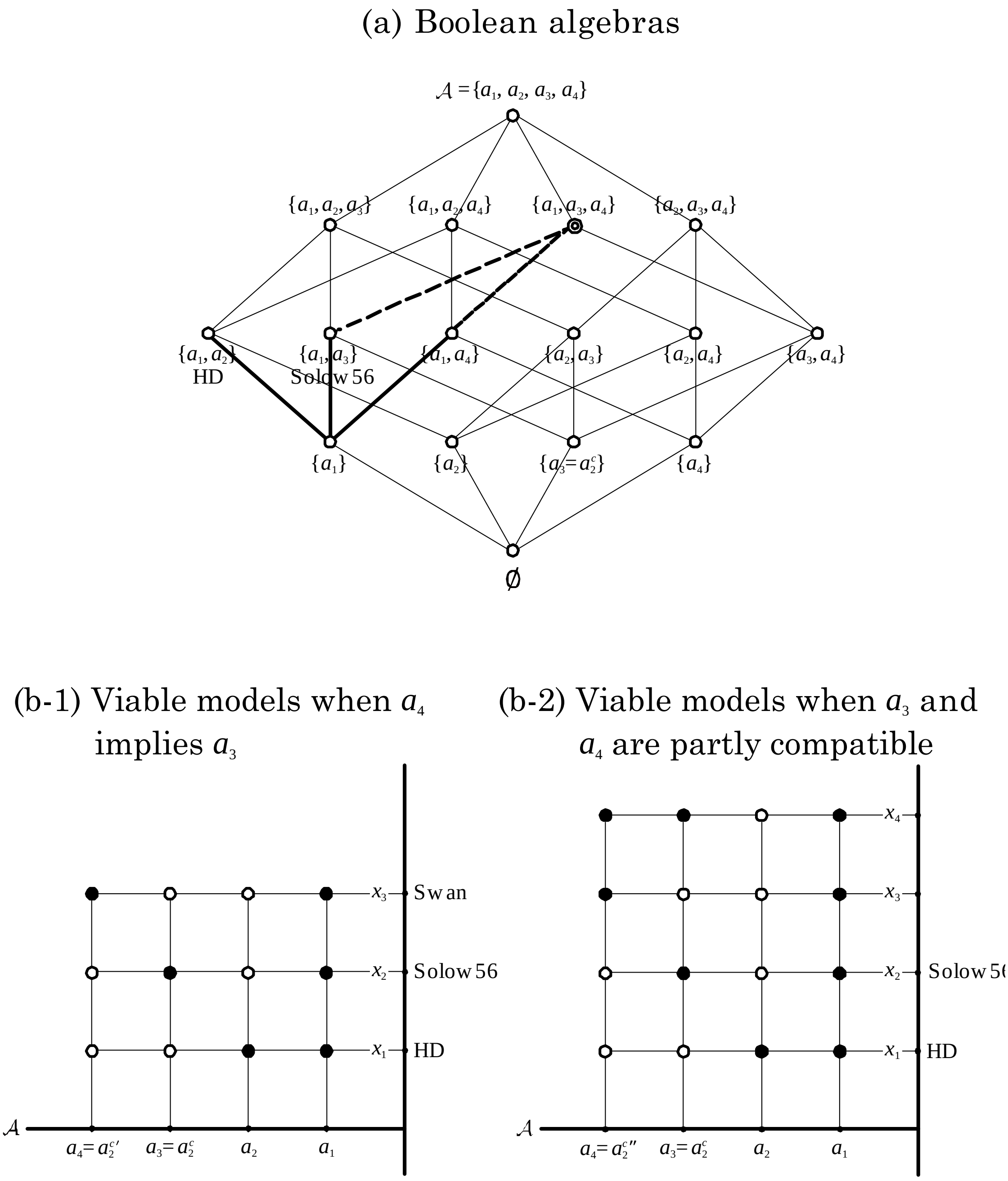
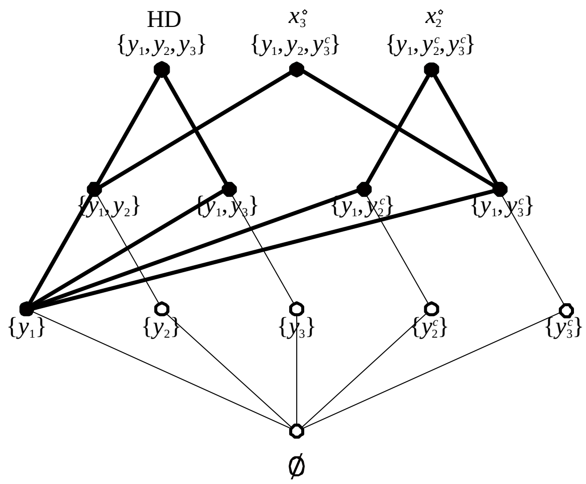


Figure 11: Boolean Algebras and Viable Models in Theory Space with $\mathcal{A} = \{a_1, a_2, a_3 = a_2^c, a_4\}$

(a) Boolean sub-algebras of viable implication-sets



(b) Viable implication-sets of model x_1 and potential models x_2° and x_3°

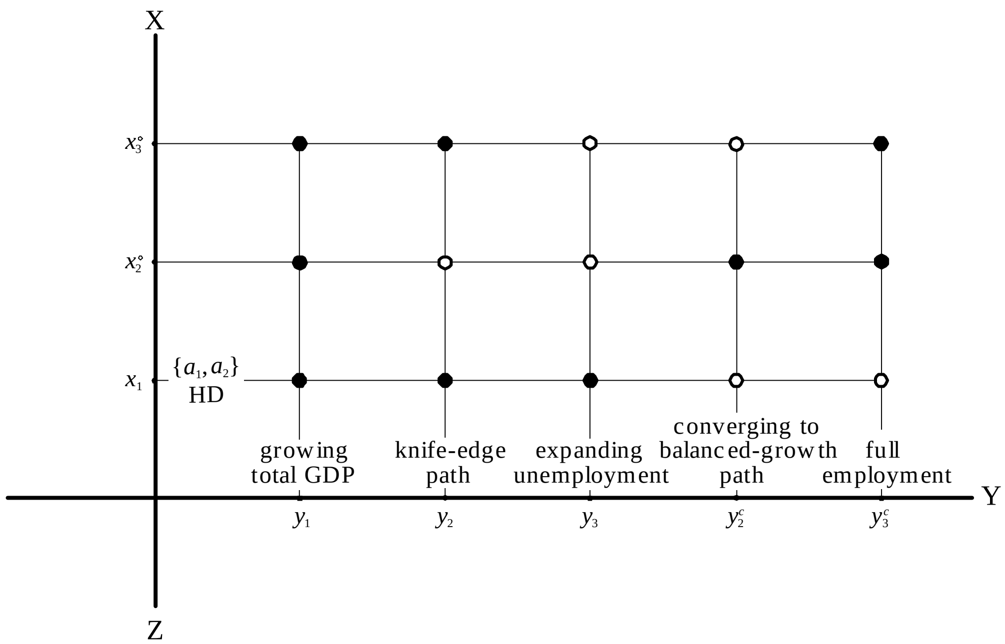


Figure 12: Boolean Algebras of Viable-Implication Sets and Potential Models in Science Space with $\mathcal{Y} = \{y_1, y_2, y_3, y_2^c, y_3^c\}$

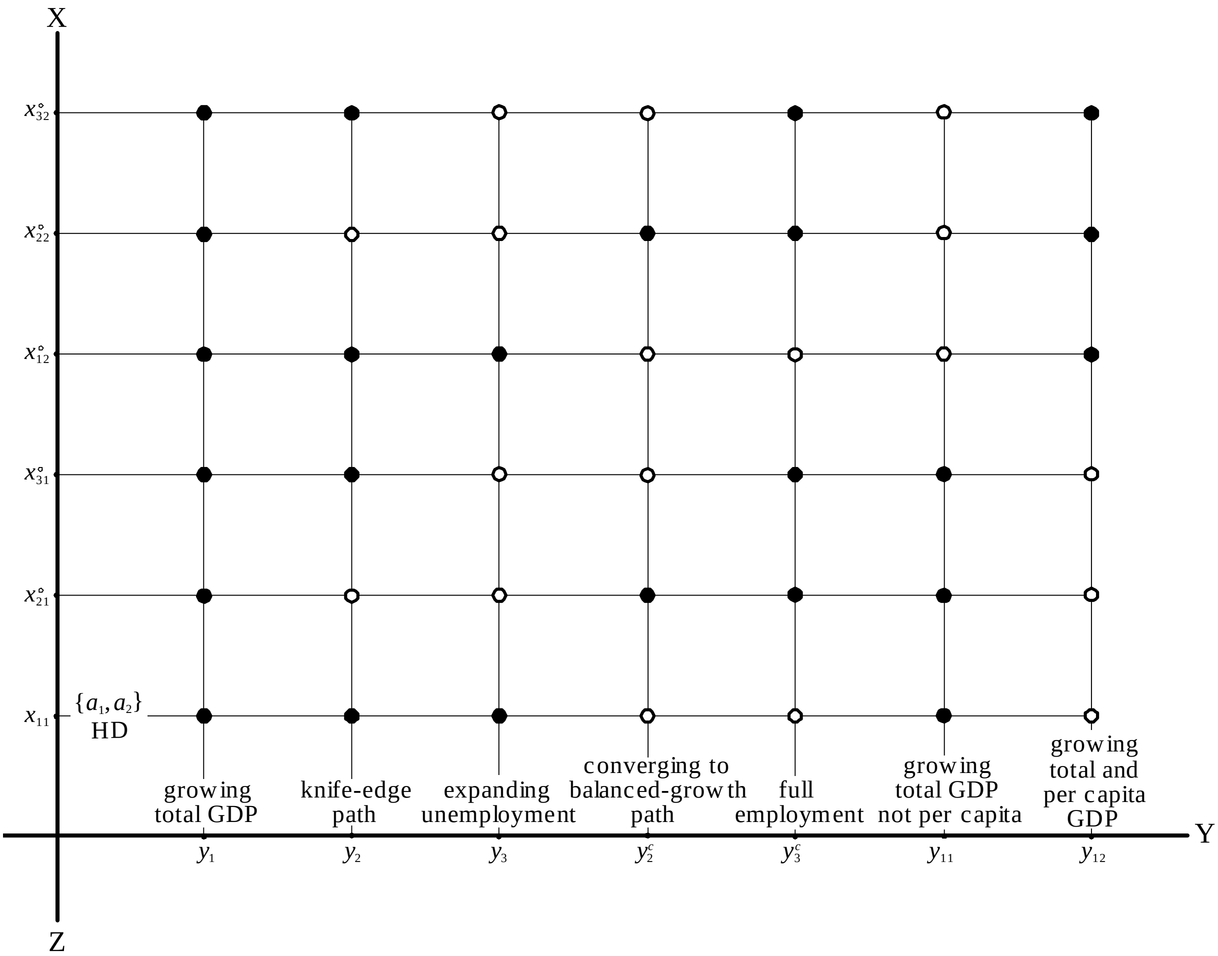


Figure 13: Viable implication-sets and potential models in Science space with $\mathcal{Y} = \{y_1, y_2, y_3, y^c_2, y^c_3, y_{11}, y_{12}\}$