Epistemic Logic with Functional Dependency Operator

Yifeng Ding

Group in Logic and the Methodology of Science, UC Berkeley. *

yf.ding@berkeley.edu

6/8/2017

Abstract

Epistemic logic with non-standard knowledge operators, especially the "knowing-value" operator, has recently gathered much attention. With the "knowing-value" operator, we can express knowledge of individual variables, but not of the relations between them in general. In this paper, we propose a new operator *Kf* to express knowledge of the functional dependencies between variables. The semantics of this *Kf* operator uses a function domain which imposes a constraint on what counts as a functional dependency relation. By adjusting this function domain, different interesting logics arise, and in this paper we axiomatize three such logics in a single agent setting. Then we show how these three logics can be unified by allowing the function domain to vary relative to different agents and possible worlds. A multiagent axiomatization is given in this case.

1 Introduction

De re knowledge or in general non-standard knowledge in epistemic logic is attracting continuing attention. This line of research started from the very beginning of epistemic logic: Hintikka discussed a "knowing-who" operator in [3], and Plaza a "knowing-value" operator *Kv* in his seminal work [4]. However, it is the recent effort in providing formal semantics and axiomatizations of those non-standard knowledge operators, as outlined in the survey [8], that layed a solid foundation for further investigation. Among all the non-standard knowledge operators axiomatized so far, the "knowing-value", or equivalently the "knowing-what" operator, has received most attention, partly due to its mathematical elegance and partly because of its potential application in information security reasoning. Recent major development of this *Kv* operator started with the axiomatization in [10, 9], followed by the

*The author would like to give special thanks to Malvin Gattinger and Wesley H. Holliday for their unreserved helpful comments.

simplification of the semantics in [2] and the enrichment of the language through announcing values and propositions in [1, 7].

Building on the above results about the "knowing-value" operator, this paper considers the knowedge of the *functional dependency* between variables, which is a natural extension of the knowledge of individual variables to the knowledge of relations among variables. The precise meaning of "knowing a/the functional dependency between variables" is not easy to pin down and might be context sensitive, as illustrated by the difficulty to choose the correct article here: it is safe to say "knowing the value of a variable" since a variable can only take one value in the actual world (or any world), but there might be quite a lot functions, different from each other, yet all governing the relation between the same two variables in a set of possible worlds. We postpone further discussion to the last section, but it should be intuitive that "functionality" is at least a minimal requirement, that is, to know any functional dependency between variables c and d, at least for any two possible worlds where c has the same value, d should also have the same value, however different from the value of c.

Here one natural choice is to make functionality the only requirement of "knowing a/the functional dependency between variables", and both [7, 1] made this choice. The key intuition behind this choice is that, what matters in the end are the values of variables. Recall how implication in Heyting algebras for intuitionistic logic is defined: $p \to q$ is the weakest proposition such that if conjoined with p by taking conjunction, we get something stronger than q, or in other words, we are able to infer q. In our knowing-value context, we might also be interested and only interested in knowing the values. Then, functional dependency of d upon c should be interpreted as the weakest proposition such that if "conjoined" with the knowledge of the value of d, we are able to infer the value of c.

The weakest proposition possessing this bridging-the-gap property depends on how we interpret the word "conjoin" here. If it is taken to be the propositional conjunction, then what we get is again the propositional implication $Kv(c) \to Kv(d)$. If "conjoin" means revealing the actual value of c to the agent, then $[c]Kv(d)$ in [7] is an exact formalization. Model-theoretically speaking this means that functionality between c and d holds on the set of possible worlds where the value of c is correct, and consequently, once all possible worlds where c's value is wrong are eliminated, the value of d becomes fixed and hence known. If "conjoin" means to entertain the hypothesis that one of the epistemically possible values of c obtains, then the functionality condition from c to d among all possible worlds is the minimal requirement. This is equivalent to $K[c]Kv(d)$, which says: I know that for all possible values that c can take, once that is revealed to be the real value of c, the value of d will also be known. In [1], this is exactly the semantics of $K^c d$.

Another famous work on dependency taking functionality as the only requirement is Dependence Logic [6, 5]. The team semantics it uses for the dependence atom $=(c, d)$ is exactly the functionality condition, though the teams in a model do not originate from an epistemic setting.

The semantics to be proposed in this paper will differ from the above pure functionality approach and will subsume it as a special case. But the key inspiration comes from the basic strategy explained in [8]: pack an existential quantifier and a modal quantifier together in the form of $\exists x \Box \phi(x)$. Under this pattern, the knowledge of the functional dependency of variables c and d is expressed as: there exists a function f in a predetermined function domain $\mathbf{F}$ which works, in the sense that $d = f(c)$, in all epistemic scenarios. Thus, **F** can be seen as an agent's prior knowledge about possible functional dependency relations, and to know the dependency between variables is to find a possible function that works or explains all possibilities. To put it more colloquially, to know the functional dependency between c and d is not simply to see that functionality holds between them, but also to see that the functional relation "make sense". Let us use $Kf(C, d)$ to express this knowledge of functional dependency of d upon a finite set of variables C.

As argued above, when "knowing-dependency" serves as a tool for expressing potential "knowing-value", we do not need a requirement stronger than functionality. But this is not always the case. Consider a typical scenario in information security: agent `A` receives an encrypted message $d = \mathtt{enc}(c)$ from agent `B`. Ideally, `A` knows the value of d, say $d = 0$, but knows nothing about c. So the epistemically possible worlds for `A` are

$$\{c = 0, d = 0\}, \{c = 1, d = 0\}, \{c = 2, d = 0\}, \cdots .$$

Certainly the functionality from c to d holds as d has only one possible value. But agent `A` is apparently ignorant about the functional relationship between variables d and c. The witness to the functionality here is the constant function 0, which is extremely unlikely to be the encryption function `enc` that `B` uses. So agent `A` would not in this case assert that she knows that the message d she receives is derived from the message c that `B` intends to send through some encryption: no encryption function she deems possible would allow all those possibilities. Thus, to claim the knowledge of the functional dependency of d on c, we do need something more than functionality. With our operator Kf, we can use $Kf_{\mathtt{A}}(\{c\}, d)$ to express "`A` knows a functional dependency relation between c and d that is plausible in the information security context", if we let **F** to be the set of all functions that is plausible in this context.

Thus, the Kf operator can be used to model scenarios where the value of variables in the realized world (the agent's world) is not the sole concern of the agent. It might be that our agent does not want an inexplicable relationship between variables, or it might be that the agent requires that any functional dependency she knows to be applicable not only to her actual world but also to worlds metaphysically possible or worlds evolved in time, where some *a priori* rules preclude too strange functional dependency relationships. In the previous case, certainly d is known to `A` already, but the constant function that witnesses the functionality there is not likely to be applicable to another round of message exchange.

In the rest of the paper, we first define the logic that incorporates knowledge K, "knowing-value"Kv and "knowing-function"Kf operators which we call **LKVF** and the corresponding base axiom system $\mathbb{LKVF}$. Then we show how different domains of functions, viewed as a parameter of **LKVF**, induce different sets of validities and axioms. Then all those cases will be put into a unified framework where a multiagent logic with the same operators is axiomatized. In the last section, we will discuss further interpretations of "knowing a/the functional dependency between variables" and possible future work.

2 Preliminaries

2.1 Syntax and Semantics of LKVF

Definition 1 (Syntax) Given a countably infinite set **P** of propositional letters and a set **Q** of the names of variables, the formulas in **LKVF** are defined by:

$$\phi ::= \top \mid p \mid Kv(d) \mid Kf(C, d) \mid (\phi \land \phi) \mid \neg\phi \mid K\phi$$

where $p \in \mathbf{P}$, $d \in \mathbf{Q}$, and $C \subseteq_{fin} \mathbf{Q}$. $\subseteq_{fin}$ means a finite subset, possibly empty.

Here $Kv(d)$ is to be interpreted as "knowing the value of d", and $K\phi$ "knowing that ϕ is the case". $Kf(C, d)$ says that the agent knows a functional dependency relationship from C to d. By convention, we set $\bot, (\phi \lor \psi), (\phi \to \psi)$ as $\neg\top, \neg(\neg\phi \land \neg\psi), \neg(\phi \land \neg\psi)$, and omit unnecessary parentheses. We also write $Kf(c, d)$ as an abbreviation of $Kf(\{c\}, d)$.

In order to interpret the Kf operator in **LKVF**, we need a predefined domain **G** of possible values for variables in **Q**, and a set of **F** functions on this **G**. **F** might contain polyadic functions in $\mathbf{G}^{\mathbf{G}^n}$ and also zero-adic functions. Formally $\mathbf{F} \subseteq \bigcup_{n=0}^{\infty} \mathbf{G}^{\mathbf{G}^n}$. It is important to note here that in this setting, **F** and **G** are important parameters of **LKVF** instead of parts of the models. In other words, they are shared by all models in the logic.

As we are considering single agent S5, no explicit accessibility relation is needed. So formally, a model is:

$$\mathcal{M} = \langle W, U, V \rangle$$

where W is the set of possible worlds, $U : W \times \mathbf{P} \to \{0, 1\}$ is the assignment for propositional letters, and $V : W \times \mathbf{Q} \to \mathbf{G}$ is the assignment for variables. For any finite subset C of **Q**, we fix an order of the elements in C and define $V(w, C) = \langle V(w, d) \mid d \in C \rangle$. When C is empty, this degenerates into the unique empty tuple. We call this the joint assignment of variables in C, and whenever we have a function from **Q** to **G**, if it is applied to a set C, we mean this joint assignment. Now the truth conditions are:

Definition 2 (Semantics)

$$\begin{array}{lcl}
\mathcal{M}, w \vDash \top & & \text{always} \\
\mathcal{M}, w \vDash p & \Leftrightarrow & U(w, p) = 1 \\
\mathcal{M}, w \vDash Kv(d) & \Leftrightarrow & \exists x \in \mathbf{G}, \forall w' \in W, V(w', d) = x \\
\mathcal{M}, w \vDash Kf(C, d) & \Leftrightarrow & \exists f \in \mathbf{F}, \forall w' \in W, V(w', d) = f(V(w', C)) \\
\mathcal{M}, w \vDash \phi \land \psi & \Leftrightarrow & \mathcal{M}, w \vDash \phi \text{ and } \mathcal{M}, w \vDash \psi \\
\mathcal{M}, w \vDash \neg\phi & \Leftrightarrow & \text{not } \mathcal{M}, w \vDash \phi \\
\mathcal{M}, w \vDash K\phi & \Leftrightarrow & \forall w' \in W, \mathcal{M}, w' \vDash \phi
\end{array}$$

Here the Kv operator has the same meaning as that of Kv in [10]: $Kv(d)$ means that under current epistemic uncertainty, the value of d is certain. The new operator $Kf(C, d)$ here means: the agent can find a function in the set of available functions **F** that can be used to explain the functional dependency relation between C and d. While both operators have the same structure in their semantics, namely $\exists\Box$, the key difference here is that, if $Kv(d)$ is true, only one value will be the witness, yet for Kf this is usually not the case.

To summarize, our logic **LKVF** extends the standard propositional epistemic logic by adding $Kv(d)$ and $Kf(C,d)$ to the language, adding a valuation of the variables to the models, and introducing a new function domain **F** as part of the logic. Now it has the following parameters:

- **P**: the set of propositional letters
- **Q**: the set of variable names
- **G**: the set of values that variables can take
- **F**: the set of functions that the agent deems possible *a priori*.

All of them will have some effect on the validities of **LKVF**, but **P** and **Q** will remain unchanged throughout the whole paper, since they can be viewed as part of the language. **G** needs to be large for completeness results, and we will specify how large it should be. **F** will change the validities in **LKVF** in an interesting way. Thus, it will be one of the main focuses of this paper. Later we show how **F** can also be put into the models.

2.2 Base Axiom System and Soundness Condition

As defined above, the Kf operator expresses functional dependencies among variables and thus resembles the dependency relation in database theory. Using Armstrong's three axioms in [11], we obtain this base system $\mathbb{LKVF}$:

`TAUT`	Propositional Tautologies		
`K`	$K(\phi \to \psi) \to (K\phi \to K\psi)$	`KV4`	$Kv(d) \to KKv(d)$
`T`	$K\phi \to \phi$	`KV5`	$\neg Kv(d) \to K\neg Kv(d)$
`4`	$K\phi \to KK\phi$	`KF4`	$Kf(C,d) \to KKf(C,d)$
`5`	$\neg K\phi \to K\neg K\phi$	`KF5`	$\neg Kf(C,d) \to K\neg Kf(C,d)$

$$\begin{array}{rl} \texttt{PROJ} & Kf(C,c) \quad c \in C \\ \texttt{TRAN} & \left(\bigwedge_{d\in D} Kf(C,d)\right) \wedge Kf(D,e) \to Kf(C,e) \\ \texttt{VF} & \left(\bigwedge_{c\in C} Kv(c)\right) \wedge Kf(C,d) \to Kv(d) \end{array}$$

$$\mathbb{MP} \quad \frac{\phi, \phi \to \psi}{\psi} \qquad \mathbb{NEC} \quad \frac{\phi}{K\phi}$$

Here only the projectivity and transitivity axioms are used. The reason is that in our language the syntax of Kf allows only one variable to be dependent upon a set of variables, not a set upon a set. Thus, the additivity property $Kf(A,B) \wedge Kf(A,C) \to Kf(A, B \cup C)$ dealing with the second set of variables after Kf is not used and will follow from the properties of the conjunction if we define $Kf(C,D)$ to be $\bigwedge_{d \in D} Kf(C,d)$. Then the augmentation axiom in the usual presentation of Armstrong's axioms follows from additivity, projectivity, and transitivity. To show this, suppose $Kf(A,B)$. By projectivity, $Kf(A \cup C, C)$ and $Kf(A \cup C, A)$. Together with the assumption $Kf(A,B)$, we have $Kf(A \cup C, B)$. So by additivity applied to $Kf(A \cup C, C)$ and $Kf(A \cup C, B)$, $Kf(A \cup C, B \cup C)$.

By convention, an empty conjunction is $\top$. So when the set D in `TRAN` is empty, it actually says $Kf(\emptyset, e) \to Kf(C,e)$ for all $C \subseteq \mathbf{Q}$. And when the set C in `VF` is empty, it says $Kf(\emptyset, d) \to Kv(d)$.

We will discuss the axiomatizations of three different settings using a large, a small, and an intermediate $\mathbf{F}$ in **LKVF** respectively. For them, we either use $\mathbb{LKVF}$ itself or add some other special axioms. To simplify repetitive work, here we give a condition on $\mathbf{F}$ in **LKVF** for the soundness of $\mathbb{LKVF}$:

Proposition 1 When $\mathbf{F}$ satisfies the following, $\mathbb{LKVF}$ is sound with respect to **LKVF**:

- For every $i, j \in \mathbb{N}$ such that $0 < i \leq j$ and function $f : \mathbf{G}^j \to \mathbf{G}$, $f(x_1, x_2, \ldots, x_j) = x_i$ is in $\mathbf{F}$. We denote this special projection function as $id_{i,j}$.
- For every $f \in \mathbf{F}$, if f is n-ary with $n \geq 1$, then for every $g_1, \ldots g_n \in \mathbf{F}$, $f(g_1(), \ldots, g_n()) \in \mathbf{F}$. Namely, $\mathbf{F}$ is closed under function composition.

Proof. Here we only prove the soundness of the three less trivial axioms:

- By the first property of $\mathbf{F}$, `PROJ` holds. If $d \in C$, suppose d appears in C as the ith variable, then $V(w,d) = V(w,C)[i]$ always holds, and thus the witness of $Kf(C,d)$ is $id_{i,|C|}$.
- By the second property of $\mathbf{F}$, `TRAN` holds. The antecedent of this axiom states the existence of f and g_is in the second property. So the composition of f and g_is exists in $\mathbf{F}$, which witnesses the consequent of `TRAN`.
- We want to show

  $$\texttt{VF} : \left(\bigwedge_{c \in C} Kv(c) \right) \wedge Kf(C,d) \to Kv(d).$$

  Let C be enumerated as $c_1, \ldots, c_n$ and suppose the antecedent in `VF` holds. Then $\bigwedge_{c \in C} Kv(c)$ is true. This means we have a tuple $\bar{a} \in \mathbf{G}^n$ such that

  $$\forall w, V(w,C) = \bar{a}.$$

  Further we have $Kf(C,d)$, which means we have a $f \in \mathbf{F}$ such that

  $$\forall w, V(w,d) = f(V(w,C)) = f(\bar{a}).$$

  Thus, there exists an element $b := f(\bar{a}) \in \mathbf{G}$ such that d evaluates to it in all possible worlds. ■

We will briefly mention how $\mathbf{F}$ is going to satisfy this soundness condition in all the following cases.

3 Full Domain of Functions

In this section, we deal with the case where $\mathbf{F}$ is as large as possible, namely $\mathbf{F} = \bigcup\{\mathbf{G}^{\mathbf{G}^i} \mid i \in \mathbf{N}\}$. Now the Kf operator degenerates into a functionality test, as all functions are allowed:

$$\mathcal{M}, w \vDash Kf(C, d) \Leftrightarrow$$
$$\forall w_1, w_2 \in W, V(w_1, C) = V(w_2, C) \Rightarrow V(w_1, d) = V(w_2, d).$$

This is true because once we have the right hand side true, we will obtain a partial function f satisfying $\forall w \in W, f(V(w, C)) = V(w, d)$. And it is trivial to extend this partial function into a total function.

Now, if $\mathcal{M}, w \vDash Kv(d)$, then $\forall w_1, w_2 \in W, V(w_1, d) = V(w_2, d)$, so the right hand side of the above truth condition holds, and consequently, $Kf(C, d)$ is true in $\mathcal{M}, w$. This justifies the soundness of our new axiom in this case:

$$\mathtt{EXT}: \quad Kv(d) \to Kf(C, d)$$

where $C \subseteq_{fin} \mathbf{Q}$, possibly empty. We name this axiom `EXT` because it means that in this case every function on $\mathbf{G}$, regardless of its meaning, can serve as a witness of the truth condition of Kf. Further, $\mathbf{F}$ satisfies the condition given in Proposition 1, so $\mathbb{LKVF}$ + `EXT` is sound. In the following, we prove that if $\mathbf{G}$ is sufficiently large, then $\mathbb{LKVF}$ + `EXT` is in fact complete as well.

Given an arbitrary set A of formulas consistent in $\mathbb{LKVF}$ + `EXT`, the Lindenbaum lemma enables us to construct a maximal consistent set Γ such that $A \subseteq \Gamma$. Now to build a model for Γ, we need to accompany this Γ by other maximal consistent sets (possible worlds). For example, if we have $\neg Kf(C, d)$ in Γ, then we need two possible worlds on which the values of C coincide while the values of d on them diverge.

To this end, we first define some useful sets. Given any maximal consistent set Γ, define

$$K_\Gamma = \{\phi \mid K\phi \in \Gamma\}, Kv_\Gamma = \{d \mid Kv(d) \in \Gamma\}.$$

They collect all the propositional and the value knowledge respectively in Γ. For any $C \subseteq \mathbf{Q}$, we say C is **closed under Kf in Γ** if for all $C_f \subseteq_{fin} C$ and $d \in \mathbf{Q}$ such that $Kf(C_f, d) \in \Gamma$, we have $d \in C$ as well. Using axioms `TRAN` and `PROJ`, it is not hard to see that for all $C \subseteq_{fin} \mathbf{Q}$,

$$C^{+\Gamma} := \{d \mid Kf(C, d) \in \Gamma\}$$

is closed under Kf in Γ and $C \subseteq C^{+\Gamma}$. This can be seen as the dependency hull of the finite set C. An important observation is that, by axiom `VF`, if $Kf(\emptyset, d) \in \Gamma$, then $Kv(d) \in \Gamma$, so $\emptyset^{+\Gamma} \subseteq Kv_\Gamma$. Also, by axiom `EXT`, if $Kv(d) \in \Gamma$ then $Kf(C, d) \in \Gamma$ for all $C \subseteq \mathbf{Q}$. So $Kv_\Gamma \subseteq C^{+\Gamma}$ for all $C \subseteq_{fin} \mathbf{Q}$, and in particular $Kv_\Gamma \subseteq \emptyset^{+\Gamma}$. So $Kv_\Gamma = \emptyset^{+\Gamma}$. This motivates us to define the set of all finitely generated closed sets:

$$M_\Gamma = \{C^{+\Gamma} \mid C \subseteq_{fin} \mathbf{Q}\}.$$

Clearly M_Γ is non-empty, and $Kv_\Gamma \in M_\Gamma$. Also, for all $X \in M$ we have $Kv_\Gamma \subseteq X$, so in other words, any finitely generated closed set contains all variables with known value. Then, we have the following disjoint decomposition of $\mathbf{Q}$ using $X \in M_\Gamma$:

$$\mathbf{Q} = Kv_\Gamma \cup (X \backslash Kv_\Gamma) \cup (\mathbf{Q} \backslash X).$$

Intuitively, the values of the variables in Kv_Γ must hold fixed among all possible worlds; the values of the variables in $X \backslash Kv_\Gamma$ must vary relative to those in Kv_Γ in a uniform way to respect the functional dependencies among them; and the values of the variables in $\mathbf{Q} \backslash X$ must vary even when all values in $X \backslash Kv_\Gamma$ are fixed, since they are not determined by X.

For example, suppose $\mathbf{Q} = \{a, b, c, d\}$, $\mathbf{G} = \mathbb{N}$, and we want to model Γ whose knowledge consists only of:

$$Kv(a), Kf(b, c)$$

and their logical consequences such as $Kf(c, a)$. Then, when considering $X = \{a, b, c\} = \{b\}^{+\Gamma}$, we have $Kv_\Gamma = \{a\}$, $X \backslash Kv_\Gamma = \{b, c\}$, and $\mathbf{Q} \backslash X = \{d\}$. Among all possible worlds, the value of a must be fixed; c must change as $c \notin Kv_\Gamma$, but it should change together with b in case of violating functionality; and d has to change even when b together with c are fixed to refute $Kf(b, d)$. Thus, one instantiation of this could be:

Kv_Γ	$a = 0$	$a = 0$	$a = 0$
$X \backslash Kv_\Gamma$	$b = 0$	$b = 1$	$b = 1$
	$c = 0$	$c = 1$	$c = 1$
$\mathbf{Q} \backslash X$	$d = 0$	$d = 1$	$d = 2$

where the columns are possible assignments. For every $X \in M_\Gamma$ which collects all closed set of variables, we need such possibilities to take care of all formulas of the form $\neg Kf(C, d)$ in Γ, because there will be one X, namely $C^{+\Gamma}$, that separates C and d. Then, the value of d can vary even when those of C are fixed.

The reason we are using only finitely generated closed subsets of $\mathbf{Q}$ is that, when $|\mathbf{Q}|$ is infinite, the cardinality remains the same. Formally, define $\mathcal{P}_f(\mathbf{Q})$ to be the collection of all finite subsets of $\mathbf{Q}$, then $|\mathcal{P}_f(\mathbf{Q})| = |\mathbf{Q}|$ when $|\mathbf{Q}| \geq \aleph_0$. Of course, when $\mathbf{Q}$ is finite, $\mathcal{P}_f$ coincides with $\mathcal{P}$, the ordinary powerset construction. Then, by the definition of M_Γ, $|M_\Gamma| \leq |\mathcal{P}_f(\mathbf{Q})|$.

Now suppose $|\mathbf{G}| \geq |\mathcal{P}_f(\mathbf{Q}) \times \{0, 1\}|$, which is the largeness condition for $\mathbf{G}$ in this case, then there exists an injection $g : M_\Gamma \times \{0, 1\} \to \mathbf{G}$. Using this g we can define a function V_p on $M_\Gamma \times \{0, 1\} \times \mathbf{Q}$ as follows:

$$V_p(\langle X, i\rangle, d) = \begin{cases} g(\emptyset, 0) & d \in Kv_\Gamma \\ g(X, 0) & d \in X \backslash Kv_\Gamma \\ g(X, i) & d \in \mathbf{Q} \backslash X. \end{cases}$$

Notice how this satisfies the informal requirement, illustrated by the example above, over the values the variables in different regions should take. When $d \in Kv_\Gamma$, its value is fixed to $g(\emptyset, 0)$. When $d \in X \backslash Kv_\Gamma$, its value depends on X as a whole but nothing else, so all variables in $X \backslash Kv_\Gamma$ change uniformly from what they are assigned by $g(\emptyset, \cdot)$. When $d \in \mathbf{Q} \backslash X$, its value further depends on i, so will change even when the values of the variables in X are fixed.

Formally, this definition allows us to show:

Proposition 2 For all $C \subseteq_{fin} \mathbf{Q}, d \in \mathbf{Q}$:

1. If $Kv(d) \in \Gamma$ then
$$\exists x \in G, \forall \langle X, i\rangle \in M \times \{0,1\}, V_p(\langle X, i\rangle, d) = x;$$
2. If $Kv(d) \notin \Gamma$ then
$$\exists \langle X, i\rangle, \langle X', i'\rangle \in M \times \{0,1\}, V_p(\langle X, i\rangle, d) \neq V_p(\langle X', i'\rangle, d);$$
3. If $Kf(C, d) \in \Gamma$ then
$$\forall \langle X, i\rangle, \langle X', i'\rangle \in M \times \{0,1\},$$
$$V_p(\langle X, i\rangle, C) = V_p(\langle X', i'\rangle, C) \Rightarrow V_p(\langle X, i\rangle, d) = V_p(\langle X', i'\rangle, d);$$
4. If $Kf(C, d) \notin \Gamma$ then
$$\exists \langle X, i\rangle, \langle X', i'\rangle \in M \times \{0,1\},$$
$$V_p(\langle X, i\rangle, C) = V_p(\langle X', i'\rangle, C) \text{ and } V_p(\langle X, i\rangle, d) \neq V_p(\langle X', i'\rangle, d).$$

Proof. For the first part, the witness is $x = g(\emptyset, 0)$ and can be verified easily. For the second part, as we observed before, $Kv_\Gamma = \emptyset^{+\Gamma} \in M_\Gamma$. Then, if $d \notin Kv_\Gamma$, on $\langle Kv_\Gamma, 0\rangle$ and $\langle Kv_\Gamma, 1\rangle$ our valuation function V_p gives different values by the injectivity of g.

For the third part, two cases are possible. If $C \subseteq Kv_\Gamma$, then $d \in Kv_\Gamma$ by VF. Then V_p assigns $g(\emptyset, 0)$ to d on all $\langle X, i\rangle$, making the consequent of the implication to be proven true throughout.

Now suppose $C \nsubseteq Kv_\Gamma$ and take $c \in C\backslash(Kv_\Gamma)$ and $\langle X, i\rangle, \langle X', i'\rangle \in M \times \{0,1\}$ such that $V_p(\langle X, i\rangle, C) = V_p(\langle X', i'\rangle, C)$. We first show $X = X'$ by focusing on this $c \notin Kv_\Gamma$. Since $c \notin Kv_\Gamma$, by the definition of V_p, there exists $j, k \in \{0,1\}$ such that
$$V_p(\langle X, i\rangle, c) = g(X, j), V_p(\langle X', i'\rangle, c) = g(X', k).$$
By the injectivity of g, they are equal only if at least $X = X'$. Based on this, if $i = i'$ then $\langle X, i\rangle = \langle X', i'\rangle$ and trivially d receives the same value from V_p.

If $i \neq i'$, recall that we assumed $V_p(\langle X, i\rangle, C) = V_p(\langle X, i'\rangle, C)$. For all $c \in C$, it follows that $c \in X$ as otherwise the values V_p gives to c differ on i and i'. Hence $C \subseteq X$ and by assumption $X \in M_\Gamma$, which means X is closed. Thus, as $Kf(C, d) \in \Gamma$, $d \in X$ as well. By definition,
$$V_p(\langle X, i\rangle, d) = g(X, 0) = g(X', 0) = V_p(\langle X', i'\rangle, d).$$

For the last part, we assume that $Kf(C, d) \notin \Gamma$. Then $d \notin C^{+\Gamma}$. By the injectivity of g and the fact that $C \subseteq C^{+\Gamma}$,
$$V_p(\langle C^{+\Gamma}, 0\rangle, d) \neq V_p(\langle C^{+\Gamma}, 1\rangle, d),$$

whereas

$$V_p(\langle C^{+\Gamma}, 0\rangle, C) = V_p(\langle C^{+\Gamma}, 1\rangle, C).$$

The above proposition handles the knowledge and ignorance about values and functional dependencies. Now we need to combine it with a traditional completeness proof for epistemic S5 logic. Denote

$$L := \{\Delta \mid \Delta \text{ is maximal consistent in } \mathbb{LKVF} + \texttt{EXT} \text{ and } K_\Gamma \subseteq \Delta\}.$$

Here L is non-empty since by axiom $\texttt{T}$, $K_\Gamma \subseteq \Gamma$ so at least $\Gamma \in L$. Then we define a model on possible worlds $W = L \times M_\Gamma \times \{0, 1\}$: $\mathcal{M} = \langle W, U, V\rangle$ where for every $\langle \Delta, C, i\rangle \in W$:

$$\begin{array}{rcl} U(\langle \Delta, C, i\rangle, p) & = & [p \in \Delta] \\ V(\langle \Delta, C, i\rangle, d) & = & V_p(\langle C, i\rangle, d) \end{array}$$

where $[p \in \Delta]$ is the indicator function of the statement $p \in \Delta$, which evaluates to 1 if the statement is true and 0 otherwise. Here each possible world has three components: a maximally consistent set which contains all formulas true at the world (truth lemma), a closed set of variables C which is responsible for instantiating the ignorance of the values of variables in C under the functional dependency constraint, and a number 0 or 1 which is responsible for instantiating the ignorance of the functionality property between variables in C and variables outside C.

Now the goal is to show a truth lemma, i.e., for all $\langle \Delta, C, i\rangle \in W, \phi \in \Delta \Leftrightarrow \mathcal{M}, \langle \Delta, C, i\rangle \vDash \phi$. To this end, we first need the following simple observation.

Proposition 3 For all $\Delta \in L$,

- $Kv(d) \in \Delta \Leftrightarrow Kv(d) \in \Gamma$
- $Kf(C, d) \in \Delta \Leftrightarrow Kf(C, d) \in \Gamma$
- $K\phi \in \Delta \Leftrightarrow K\phi \in \Gamma$.

Proof. Simply use the axioms $\texttt{4}$, $\texttt{5}$. For example, the third property follows from

$$\begin{array}{rcll} K\phi \in \Gamma & \Rightarrow & KK\phi \in \Gamma & [\text{axiom } \texttt{4}] \\ & \Rightarrow & K\phi \in K_\Gamma & [\text{definiton of } K_\Gamma] \\ & \Rightarrow & K\phi \in \Delta & [K_\Gamma \subseteq \Delta] \end{array}$$

$$\begin{array}{rcll} K\phi \notin \Gamma & \Rightarrow & \neg K\phi \in \Gamma & [\Gamma \text{ is maximally consistent}] \\ & \Rightarrow & K\neg K\phi \in \Gamma & [\text{axiom } \texttt{5}] \\ & \Rightarrow & \neg K\phi \in K_\Gamma & [K_\Gamma \text{ definition}] \\ & \Rightarrow & \neg K\phi \in \Delta & [K_\Gamma \subseteq \Delta] \\ & \Rightarrow & K\phi \notin \Delta & [\Delta \text{ is maximally consistent}]. \end{array}$$

Proposition 4 If $K\phi \notin \Gamma$, then there exists $\Delta \in L$ such that $\neg\phi \in \Delta$.

Proof. A standard exercise using necessitation and axiom $\mathtt{K}$. ■

Now we can prove the truth lemma:

Lemma 1 For all $\langle \Delta, C, i\rangle \in W, \phi \in \mathbf{LKVF}$, $\phi \in \Delta \Leftrightarrow \mathcal{M}, \langle \Delta, C, i\rangle \vDash \phi$.

Proof. By induction on ϕ, with the following possibilities:

- ϕ is a propositional letter or a boolean combination. This is standard.
- $\phi = Kv(d)$. Since $\Delta \in L$, by Proposition 3, $Kv(d) \in \Delta \Leftrightarrow Kv(d) \in \Gamma$. By Proposition 2, if $Kv(d) \in \Gamma$ then

  $$V(\langle \Theta, D, j\rangle, d) = V_p(\langle D, j\rangle, d) = V_p(\langle D', j'\rangle, d) = V(\langle \Theta', D', j'\rangle, d)$$

  for all $\langle \Theta, D, j\rangle, \langle \Theta', D', j'\rangle \in W$. If $Kv(d) \notin \Gamma$, by Proposition 2 again, there exists $\langle D, j\rangle, \langle D', j'\rangle \in M \times \{0, 1\}$ such that

  $$V(\langle \Gamma, D, j\rangle, d) = V_p(\langle D, j\rangle, d) \neq V_p(\langle D', j'\rangle, d) = V(\langle \Gamma, D', j'\rangle, d).$$

  As such,
  $$Kv(d) \in \Delta \Leftrightarrow Kv(d) \in \Gamma \Leftrightarrow \mathcal{M}, \langle \Delta, C, i\rangle \vDash \phi.$$

- $\phi = Kf(D, d)$. Similar to the last one. By Proposition 3, $Kf(D, d) \in \Delta \Leftrightarrow Kf(D, d) \in \Gamma$. By Proposition 2, $Kf(D, d) \in \Gamma \Leftrightarrow \mathcal{M}, \langle \Delta, C, i\rangle \vDash Kf(D, d)$.
- $\phi = K\psi$. By Proposition 3, $K\psi \in \Delta \Leftrightarrow K\psi \in \Gamma$. If $K\psi \in \Gamma$, then $\psi \in K_\Gamma$, so for all $\langle \Theta, D, j\rangle \in W$, as $\Theta \in L$, $\psi \in \Theta$. By the induction hypothesis, $\mathcal{M}, \langle \Theta, D, j\rangle \vDash \psi$. Thus, $\mathcal{M}, \langle \Delta, C, i\rangle \vDash K\psi$.

  On the other hand, if $K\psi \notin \Gamma$, by Proposition 4, there exists $\Theta \in M$ such that $\neg\phi \in \Theta$. By the induction hypothesis, $\mathcal{M}, \langle \Theta, \emptyset, 0\rangle \vDash \neg\psi$. So $\mathcal{M}, \langle \Delta, C, i\rangle \nvDash K\psi$. To sum up, $K\psi \in \Gamma \Leftrightarrow \mathcal{M}, \langle \Delta, C, i\rangle \vDash K\psi$. ■

From this proposition, we know that for all $\phi \in \Gamma$, $\mathcal{M}, \langle \Gamma, \emptyset, 0\rangle \vDash \phi$. As the consistent set A we chose at the very beginning is contained in Γ, $\mathcal{M}, \langle \Gamma, \emptyset, 0\rangle \vDash A$, which brings us:

Theorem 1 Given $|\mathbf{G}| \geq |\mathcal{P}_f(\mathbf{Q}) \times \{0, 1\}|$ and $\mathbf{F} = \bigcup\{\mathbf{G}^{\mathbf{G}^i} \mid i \in \mathbf{N}\}$, $\mathbb{LKVF}$+$\mathtt{EXT}$ axiomatizes **LKVF**.

4 Minimal Function Domain

In Proposition 1 we proved the soundness condition for $\mathbb{LKVF}$. Notice that the minimal function domain that satisfies this soundness condition is

$$\mathbf{F} = \{id_{i,j} \mid i, j \in \mathbb{N}, 0 < i \leq j\}.$$

In this section, we consider the axiomatization of the validities of **LKVF** with this **F**. Here, two axioms besides our base system $\mathbb{LKVF}$ are valid:

$$\begin{array}{ll} \texttt{CHOO} & Kf(C,d) \to \bigvee_{c \in C} Kf(c,d), \\ \texttt{EQU} & Kf(c,d) \to Kf(d,c). \end{array}$$

The validity of the first axiom is justified by:

$$d = id_{i,j}(c_1, c_2, \dots, c_j) = c_i = id_{1,1}(c_i),$$

and notice that when $C = \emptyset$, it degenerates to $Kf(\emptyset, d) \to \bot$ or equivalently $\neg Kf(\emptyset, d)$, which is true because no zero-ary function exists in **F**. This also means that `EXT` is unsound in this case, because even if $Kv(d)$ is true, $Kf(\emptyset, d)$ is false regardless. So $Kv(d) \to Kf(C,d)$ is in general false.

The validity of the second axiom follows from

$$d = id_{1,1}(c) = c \quad \Rightarrow \quad c = d = id_{1,1}(d).$$

Thus, $\mathbb{LKVF}$ + `CHOO` + `EQU` is sound. Given these two axioms and the fact that **F** consists only of projection functions, $Kf(c,d)$ is actually talking about the equality of c, d over all possible worlds, even though the value might not be known. This motivates the construction of the equivalence relation by $Kf(c,d)$ used below.

Now we turn to the proof of the completeness of $\mathbb{LKVF}$ + `CHOO` + `EQU`. Again, given a consistent set A, our plan is that we first extend it to a maximal consistent set Γ, then deal with its *de re* knowledge and propositional knowledge separately, and finally take their Cartesian product to obtain a model of Γ.

First, we partition **Q** into equivalence classes with equivalence relation $\sim$ defined by

$$c \sim d \Leftrightarrow Kf(c,d) \in \Gamma.$$

Its reflexivity, symmetry and transitivity follow from the axioms `PROJ`, `EQU`, and `TRAN`. Indeed, if we use the $C^{+\Gamma}$ and M_Γ construction, M_Γ will contain precisely those partitions and their unions. Every maximally consistent set, or a "world", naturally gives rise to such an equivalence relation on **Q**.

For every $c \in \mathbf{Q}$, define $[c] = \{d \mid c \sim d\}$, and for every $C \subseteq \mathbf{Q}$, define $[C] = \langle [c] \mid c \in C \rangle$, the collection of the equivalence classes which contain at least one of its elements. In particular, $[Kv_\Gamma] = \langle [c] \mid Kv(c) \in \Gamma \rangle$.

Now, if $|\mathbf{G}| \geq |\mathbf{Q}| \geq |[\mathbf{Q}]|$, then there will be two injections from $[\mathbf{Q}]$ to **G**, u and v, such that

$$u([c]) = v([c]) \Leftrightarrow [c] \in [Kv_\Gamma].$$

For example, we can let u be any injection and then make a rotation over the function values of u on $[\mathbf{Q}] \backslash [Kv_\Gamma]$ to obtain v in case of **Q** being finite, or let $v(d)$ be the successor of $u(d)$ for $d \in [\mathbf{Q}] \backslash [Kv_\Gamma]$ in case of **Q** being infinite (assuming it can be well ordered). We do not need to seek more valuations of variables to prove the truth lemma in this case or to instantiate the ignorances of the knowledge about values in Γ. Any one of them is capable of refuting $Kf(C, d) \notin \Gamma$ and together they instantiate $Kv(d) \notin \Gamma$.

Defining V_p as a function from $\{u, v\} \times \mathbf{Q}$ to $\mathbf{G}$ by $V_p(t, d) = t([d])$, the following proposition is true:

Proposition 5 For any $d \in \mathbf{Q}, C \subseteq_{fin} \mathbf{Q}$:

1. if $Kv(d) \in \Gamma, \exists x \in G, \forall t \in \{u, v\}, V_p(t, d) = x$
2. if $Kv(d) \notin \Gamma, \exists t, t' \in \{u, v\}, V_p(t, d) \neq V_p(t', d)$
3. if $Kf(C, d) \in \Gamma, \exists f \in \mathbf{F}, \forall t \in \{u, v\}, f(V_p(t, C)) = V_p(t, d)$
4. if $Kf(C, d) \notin \Gamma, \forall f \in \mathbf{F}, \exists t \in \{u, v\}, f(V_p(t, C)) \neq V_p(t, d)$.

Proof. The first two parts are immediate from the definition of u, v: $Kv(d) \in \Gamma \Leftrightarrow [d] \in [Kv_\Gamma] \Leftrightarrow u([d]) = v([d]) \Leftrightarrow V_p(u, d) = V_p(v, d)$.

For the third property, suppose $Kf(C, d) \in \Gamma$ and enumerate C by $c_1, \ldots, c_j$. By axiom CHOO and the maximality of Γ, there exists i such that $Kf(c_i, d) \in \Gamma$ and thus $[d] = [c_i]$. Now, for every $t \in \{u, v\}, V_p(t, C) = \langle [c_1], [c_2], \ldots, [c_j] \rangle$, so $[d] = id_{i,j}(V_p(t, C))$ and we see that the functional relation between C, d is $id_{i,j}$.

For the last one, suppose $Kf(C, d) \notin \Gamma$. It follows that $[d] \notin [C]$ because otherwise, $[d] \in [C]$ and there exists $c \in C, [d] = [c]$, hence $Kf(c, d) \in \Gamma$. By axiom PROJ, $Kf(C, c) \in \Gamma$, and then by axiom TRAN, $Kf(C, d) \in \Gamma$, which contradicts the assumption. Again enumerate $C = \langle c_1, \ldots, c_j \rangle$. Since u is injective and $[d] \notin [C]$, for all $c_i \in C$, $u([d]) \neq u([c_i])$. Thus, for every j-ary function $id_{i,j} \in \mathbf{F}$, $id_{i,j}(V_p(u, C)) = u([c_i]) \neq u([d])$. Actually we can use v here as well. The reason we need both of them is that we need to instantiate $\neg Kv(d)$ for $d \notin Kv_\Gamma$. ∎

To build a model for Γ, define

$$
\begin{aligned}
L &= \{\Delta \mid \Delta \text{ is a maximal consistent set}, K_\Gamma \subseteq \Delta\} \\
W &= L \times \{u, v\} \\
U(\langle X, t \rangle, p) &= [p \in X] \\
V(\langle X, t \rangle, d) &= V_p(t, d) \\
\mathcal{M} &= \langle W, U, V \rangle.
\end{aligned}
$$

Then we have the following truth lemma:

Lemma 2 For all $\langle \Gamma, t \rangle \in W$, $\langle \Gamma, t \rangle \vDash \phi$ if and only if $\phi \in \Gamma$.

Proof. The proof is similar to that of Lemma 1. The difference is that we need to use Proposition 5 instead of Proposition 2. ∎

The completeness of **LKVF** + CHOO + EQU follows, so we conclude:

Theorem 2 Given $|\mathbf{G}| \geq |\mathbf{Q}|, \mathbf{F} = \langle id_{i,j} \mid i, j \in \mathbb{N}, 0 < i \leq j \rangle$, $\mathbb{LKVF}$+ axiomatizes **LKVF**.

5 Intermediate Function Domain

In the previous two sections, we considered the minimal and the maximal function domains subject to our soundness condition. As we can see, in both cases the axiomatizations require some axioms besides the base system $\mathbb{LKVF}$. And those axioms are not very intuitive if we intend to interpret *Kf* as "knowing a/the functional dependency". In this section, we show that we can construct a function domain such that if $\mathbf{F}$ is set to it, $\mathbb{LKVF}$ will be complete and no extra axiom is needed. The construction is somewhat artificial but in the next section, we can view this as just one step of a completeness proof at a higher level.

The main difficulty here is to refute the axiom scheme `EXT` used in the axiomatization of the full function domain case. `EXT` is validated in that case because whenever the value of a variable is known, a constant function can be used to explain the functional dependency between it and any other variables in all epistemic possibilities. Thus, to refute this scheme as an axiom, we must make sure that the function domain encodes information more than just functionality so that we can refute $Kf(c, d)$ even when functionality holds, such as when $Kv(d)$ is true. The function domain to be constructed below will enable a suitably constructed model to refute $Kf(C, d)$ without ever looking into the functionality condition.

To do this, we go to higher dimensions by assuming $\mathbf{G} = 2^{\mathcal{P}_f(\mathbf{Q})}$, interpreted as functions from the finite subsets of $\mathbf{Q}$ to $\{0, 1\}$ or as a rather long sequence indexed by $\mathcal{P}_f(\mathbf{Q})$ where at each index (dimension) C we can choose from $\{0, 1\}$. This is actually only a size requirement, since so long as $|\mathbf{G}| \geq |2^{\mathcal{P}_f(\mathbf{Q})}|$, we can always embed $2^{\mathcal{P}_f(\mathbf{Q})}$ into $\mathbf{G}$ by an injection. For any $x \in \mathbf{G}$ and $C \subseteq_{fin} \mathbf{Q}$, we use $x[C]$ to retrieve the image of C under x, which will be 0 or 1. Now we construct the intermediate $\mathbf{F}$:

Definition 3 Let $\mathbf{F}$ be the collection of the functions f satisfying the following constraints: where y is $f(x_1, \ldots x_n)$, for all $C \subseteq_{fin} \mathbf{Q}$,

$$x_1[C] = x_2[C] = \ldots = x_n[C] = 0 \quad \Rightarrow \quad y[C] = 0.$$

Alternatively, where

$$\mathrm{Lmax}_n = \{f \in \mathbf{G}^{\mathbf{G}^n} \mid \forall C \subseteq_{fin} \mathbf{Q}, f(x_1, \ldots x_n)[C] \leq \max(x_1[C], \ldots x_n[C])\},$$

with $\max() = 0$, define $\mathbf{F} = \bigcup_{i \in \mathbb{N}} \mathrm{Lmax}_i$.

Notice that the requirement is specified for all dimensions individually, and they do not interfere with each other. This allows us to do constructions and proofs for each dimension separately.

Now we can check that this $\mathbf{F}$ satisfies the soundness condition. Projection functions are all included in $\mathbf{F}$ because they all satisfy the above constraint: for any $C \subseteq_{fin} \mathbf{Q}$, either $x_i[C] = y[C] = 1$, where the antecedent and the consequent are both false, or $x_i[C] = y[C] = 0$, where they are both true. For compositionality, let $h = f(g_1, \ldots g_n)$. If all inputs to h are 0 at any dimension C, then since $g_1, \ldots g_n \in \mathbf{F}$, they evaluate to 0 at dimension C. Then all inputs to f are 0 at this dimension C. So as $f \in \mathbf{F}$, it evaluates to 0 as well. Thus, h is in $\mathbf{F}$.

To prove the completeness of $\mathbb{LKVF}$ with respect to **LKVF** with this new function domain **F**, again the satisfiability of any maximal consistent set Γ is required, and the crucial step is still the construction of a set of valuations such that the formulas of the form $Kv(d), \neg Kv(d), Kf(C, d)$, and $\neg Kf(C, d)$ in Γ are satisfied. Indeed, for this purpose, we only need two valuations, a situation similar to that in the case of the minimal function domain. This is because when $\neg Kf(C, d) \in \Gamma$, we are refuting $Kf(C, d)$ not by a failure of functionality but by a failure of conformation to **F**. Breaking functionality requires at least two possible value assignments, but if **F** says no, a single possibility is too many. Recall the $C^{+\Gamma}$ we used in the previous two cases, which is defined as $\{d \in \mathbf{Q} \mid Kf(C, d) \in \Gamma\}$. Now we need to define a slightly different M_Γ:

$$\{C^{+\Gamma} \mid C \subseteq_{fin} \mathbf{Q}\} \cup \{Kv_\Gamma\}.$$

This is the collection of all finitely generated closed sets plus Kv_Γ. We need this extra union since axiom `EXT` is not available now, which means Kv_Γ is not automatically contained in any $C^{+\Gamma}$, and it is quite possible that Kv_Γ is not finitely generated. But still, M_Γ has a cardinality no larger than $\mathcal{P}_f(\mathbf{Q})$, since if $\mathbf{Q}$ is finite, $\mathcal{P}_f(\mathbf{Q})$ contains all subsets of $\mathbf{Q}$, and if infinite, $\mathcal{P}_f(\mathbf{Q})$ is also infinite and adding one more element into it does not increase its cardinality. Thus, there is still a surjection g from $\mathcal{P}_f(\mathbf{Q})$ to M_Γ. We can think of this g as a pseudo $(\cdot)^{+\Gamma}$ function, and it does not matter which surjection we use for g. Now we can specify the two valuations we need:

Definition 4 Let g be any surjection from $\mathcal{P}_f(\mathbf{Q})$ to M_Γ. Define $V_0, V_1 : \mathcal{P}_f(\mathbf{Q}) \to \mathbf{G}$ such that for all $d \in \mathbf{Q}, C \subseteq_{fin} \mathbf{Q}$,

$$V_0(d)[C] = \begin{cases} 0 & \text{if } d \in g(C) \\ 1 & \text{if } d \notin g(C), \end{cases}$$

$$V_1(d)[C] = \begin{cases} V_0(d)[C] & \text{if } g(C) \neq Kv_\Gamma \\ 0 & \text{if } g(C) = Kv_\Gamma. \end{cases}$$

The use of V_0 is to refute $Kf(C, d)$ if $\neg Kf(C, d) \in \Gamma$, and the use of V_1 is to refute Kv if $\neg Kv(d) \in \Gamma$. Now we prove this in detail:

Proposition 6 If $Kf(C, d) \in \Gamma$, then there exists $f \in \mathbf{F}$ such that for $i \in \{0, 1\}$, $f(V_i(C)) = v_i(d)$. If $\neg Kf(C, d) \in \Gamma$, then for all $f \in \mathbf{F}$, $f(V_0(C)) \neq V_0(d)$.

Proof. To prove the first claim, assume $Kf(C, d) \in \Gamma$ with C enumerated by $c_1, \ldots c_n$. We will construct a function $f \in \mathbf{F}$ that works in both V_0 and V_1: for all $D \subseteq_{fin} \mathbf{Q}$, $V_0(d)[D] = f(V_0(C))[D]$ and $V_1(d)[D] = f(V_1(C))[D]$. Obviously this construction should be done dimension by dimension. For any $D \subseteq_{fin} \mathbf{Q}$, the possibilities are:

- $d \in g(D)$. Thus, by definition, $V_0(d)[D] = 0$. $V_1(d)[D] = 0$ as well since the only change happens when $D = Kv_\Gamma$, and even in that case, only 1 turns to 0 and not vice versa. So we can define $f(x_1, \ldots x_n)[D] = 0$. Then $V_0(d)[D] = f(V_0(C))[D]$ and $V_1(d)[D] = f(V_1(C))[D]$, regardless of what $V_0(C)$ and $V_1(C)$ are.

- $d \notin g(D)$. Since $g(D)$ is closed and $Kf(C,d) \in \Gamma$, $C \not\subseteq g(D)$. Find $c_p \notin g(D)$. Define $f(x_1, \dots x_n)[D] = x_p[D]$. This definition satisfies the requirement of $\mathbf{F}$. And it works for V_0 because $v_0(d)[D] = V_0(c_p)[D] = 1$ (both d, c_p are outside $g(D)$). It also works for V_1 because their values change to 0 together if $g(D) = Kv_\Gamma$.

To prove the second claim, recall that $C^{+\Gamma} = \{d \mid Kf(C,d) \in \Gamma\}$ is closed under Kf in Γ and contains C by axioms `TRAN` and `PROJ`. Now since $Kf(C,d) \notin \Gamma$, $d \notin C^{+\Gamma}$. As g is a surjection from $\mathcal{P}_f(\mathbf{Q})$ to M_Γ, there exists $D \subseteq_{fin} \mathbf{Q}$ such that $g(D) = C^{+\Gamma}$. Thus, by the definition of V_0, $V_0(d)[D] = 1$, while for all $c \in C \subseteq C^{+\Gamma} = g(D)$, $V_0(c)[D] = 0$. Hence $V_0(d)[D] > \max(V_0(C)[D])$, which makes it impossible to find a function $f \in \mathbf{F}$ such that $f(V_0(C)) = V_0(d)$. ■

Proposition 7 If $Kv(d) \in \Gamma$, then $V_0(d) = V_1(d)$. If $Kv(d) \notin \Gamma$, then $V_0(d) \neq V_1(d)$.

Proof. If $Kv(d) \in \Gamma$, then $d \in Kv_\Gamma$. Now for any $C \subseteq_{fin} \mathbf{Q}$, if $g(C) \neq Kv_\Gamma$, then $V_1(d)[C] = V_0(d)[C]$ by definition. If $g(C) = Kv_\Gamma$, $V_1(d)[C] = 0$, but $V_0(d)[C] = 0$ as well since $d \in Kv_\Gamma$. Thus, $V_0(d) = V_1(d)$.

If $Kv(d) \notin \Gamma$, $d \notin Kv_\Gamma$. Since we explicitly added Kv_Γ to Γ, $Kv_\Gamma \in M_\Gamma$, and we can find a $C \subseteq_{fin} \mathbf{Q}$ such that $g(C) = Kv_\Gamma$. Then, using the definition of V_0 and V_1, we know $V_0(d)[C] = 1$ but $V_1(d)[C] = 0$, because $g(C) = Kv_\Gamma$ and we assumed $d \notin Kv_\Gamma$. Thus, $V_1(d) \neq V_0(d)$. ■

Based on the previous two propositions, we can build a model for Γ by defining

$$
\begin{array}{rcl}
L & = & \{\Delta \mid \Delta \text{ is a maximal consistent set, } K_\Gamma \subseteq \Delta\} \\
W & = & L \times \{0,1\} \\
U(\langle X, t\rangle, p) & = & [p \in X] \\
V(\langle X, t\rangle, d) & = & V_t(d) \\
\mathcal{M} & = & \langle W, U, V\rangle.
\end{array}
$$

With a proof which is essentially the same as the proof of the truth lemma Lemma 1 in the full function domain case, using Propositions 6 and 7 instead of Proposition 2, we have:

Lemma 3 For all $\langle \Gamma, t\rangle \in W$, $\mathcal{M}, \langle \Gamma, t\rangle \vDash \phi$ if and only if $\phi \in \Gamma$.

$\mathcal{M}, \langle \Gamma, 0\rangle \vDash \Gamma$ follows from this truth lemma. This finishes the completeness proof of the intermediate case, so we have:

Theorem 3 Given $|\mathbf{G}| \geq |2^{\mathcal{P}_f(\mathbf{Q})}|$, $\mathbf{F} = \bigcup_{i \in \mathbb{N}} \mathrm{Lmax}_i$, $\mathbb{LKVF}$ axiomatizes **LKVF**.

Table 1: Choice of the function domain in **LKVF** and corresponding axiomatization

	Full	Minimal	Intermediate
$\mathbf{F} =$	$\bigcup_{i \in \mathbb{N}} \mathbf{G}^{\mathbf{G}^i}$	$\{id_{i,j} \mid i, j \in \mathbb{N}, 0 < i \leq j\}$	$\bigcup_{i \in \mathbb{N}} \mathrm{Lmax}_i$
$\lvert\mathbf{G}\rvert \geq$	$\lvert\mathcal{P}_f(\mathbf{Q}) \times \{0, 1\}\rvert$	$\lvert\mathbf{Q}\rvert$	$\lvert 2^{\mathcal{P}_f(\mathbf{Q})}\rvert$
Axiomatization	$\mathbb{LKVF}$ + `EXT`	$\mathbb{LKVF}$ + `CHOO` + `EQU`	$\mathbb{LKVF}$

6 Unifying Logic

In all the previous settings, our logic **LKVF** takes a function domain **F** as a parameter. This function domain is meant to be the set of *a priori* possible functions for functional dependencies over variables. But if this set of *a priori* possibilities is relative to the agents in discussion, then this set of functions should be variable over models instead of being part of the logic and fixed for all models. After all, an agent might hold different prior knowledge in different worlds. Also, the function domain constructed in the intermediate case is, while not nonsensical for its interesting $\leq$ max structure, still somewhat artificial for its large dimension. If this function domain is part of the model, it is at the choice of the agent under discussion.

Indeed, if we put the function domain inside the definition of a model by setting

$$\mathcal{M} = \langle \mathbf{F}, W, U, V \rangle,$$

where $\mathbf{F} : \mathbf{G} \to \mathbf{G}$ satisfies the soundness condition that it contains all projection functions and is closed under function composition, W is a set of possible worlds, U is an assignment function for propositional letters, and V is an assignment function for variables, and we leave the semantics untouched, then the soundness and completeness of $\mathbb{LKVF}$ follow immediately from the results presented so far. Using **LKVF*** to denote the logic induced by the definition of the models above, we have:

Theorem 4 $\mathbb{LKVF}$ is sound and complete with respect to **LKVF*** when $\lvert\mathbf{G}\rvert \geq \lvert 2^{\mathcal{P}_f(\mathbf{Q})}\rvert$.

Proof. Because for every model of **LKVF***, its function domain satisfies the soundness condition Proposition 1, $\mathbb{LKVF}$ is sound in all the models of **LKVF***. This shows the soundness.

For any set Γ maximally consistent with respect to $\mathbb{LKVF}$, take the **F** and the model $\mathcal{M}$ constructed in the intermediate function domain case. Then $\langle \mathbf{F}, \mathcal{M} \rangle \vDash \Gamma$ and $\langle \mathbf{F}, \mathcal{M} \rangle$ is a model of **LKVF***. Thus, every maximal consistent set is satisfiable. ■

The proof above is a direct adaptation of the completeness result in the intermediate function domain case. In that case, we built a function domain that works for all maximal consistent sets in the sense that for all maximal consistent sets Γ, this same function domain can be used to refute $Kf(C, d) \notin \Gamma$ when functionality cannot be used. This is actually the reason why the cardinality requirement for **G** is very high there. However, in the current setting where function domains are part of the models, the only thing needed is a method to build a function domain for each maximal consistent set Γ so that the functional dependency relation between C, d is rejected if $\neg Kf(C, d) \in \Gamma$. The difference will be made more clear in the following multiagent case.

6.1 Multiagent logic with variable function domain

Given an index set $\mathbf{A}$ of agents, to accommodate multiple agents, the language is now expanded to

$$\phi ::= \top \mid p \mid Kv_i(d) \mid Kf_i(C, d) \mid \neg\phi \mid (\phi \wedge \phi) \mid K_i\phi,$$

with $p \in \mathbf{P}, i \in \mathbf{A}, d \in \mathbf{G}$, and $C \subseteq_{fin} \mathbf{G}$. The only difference from the single agent language defined in Definition 1 is that now we have for each agent i a separate Kv_i, Kf_i, and K_i.

For semantics, a model is now defined as:

$$\mathcal{M} = \langle W, \langle \sim_i \rangle_{i \in \mathbf{A}}, U, V, \langle \mathbf{F}_i \rangle_{i \in \mathbf{A}} \rangle$$

where $\mathbf{F}_i$ is intended to assign a collection of functional relationships that agent i deems possible *a priori* to all possible worlds in W. Thus, for all $w \in W, i \in \mathbf{A}, \mathbf{F}_i(w)$ is required to include all projection functions and to be closed under function composition. $\sim_i$ is the epistemic accessibility relation of agent i and is required to be an equivalence relation on W, the set of possible worlds (complete epistemic scenarios). Now since $\mathbf{F}_i$ is supposed to be "prior knowledge", it is also required that if $w \sim_i w'$, then $\mathbf{F}_i(w) = \mathbf{F}_i(w')$. However, we are not assuming that the prior knowledge of any agent is public to other agents, so it is quite possible that $\mathbf{F}_j(w) \neq \mathbf{F}_j(w')$ if $j \neq i$, even when $w \sim_i w'$. In a nutshell, $\mathbf{F}_i$s are not common knowledge.

The semantic clauses are defined similarly with agent indices for knowledge sentences:

$$\begin{array}{rcl}
\mathcal{M}, w \vDash Kv_i(d) & \Leftrightarrow & \exists x \in \mathbf{G}, \forall w' \sim_i w, V(w', d) = x \\
\mathcal{M}, w \vDash Kf(C, d) & \Leftrightarrow & \exists f \in \mathbf{F}_i(w), \forall w' \sim_i w', V(w', d) = f[V(w', C)] \\
\mathcal{M}, w \vDash K\phi & \Leftrightarrow & \forall w' \sim_i w' \Rightarrow \mathcal{M}, w' \vDash \phi.
\end{array}$$

Let $\mathbf{LKVF}^*_m$ name this multiagent logic. Also, let $\mathbb{LKVF}_m$ denote the axiom system adapted from $\mathbb{LKVF}$ with indexed version of those axioms involving knowledge operators. In particular, no interaction between agents is allowed, as there are no axioms saying that we can derive any knowledge about other agents from any agent. We will see that this is precisely because we allow each agent to possess its own prior knowledge about possible functional dependencies, not necessarily known to other agents. Once we assume that $\mathbf{F}_i$s are common knowledge, interactions will arise, and we will discuss this point in the last section.

The soundness of $\mathbb{LKVF}_m$ with respect to $\mathbf{LKVF}^*_m$ follows from an indexed version of Proposition 1. For completeness we need a new construction:

Definition 5 (Dependency lattice) Given a maximal consistent set Γ in $\mathbf{LKVF}_m$ and an agent index

i, first define the indexed version of the $(\cdot)^{+\Gamma}$ operator, Cl_i^Γ, on finite subsets of $\mathbf{Q}$ as

$$\mathrm{Cl}_i^\Gamma(C) = \{d \mid \mathit{Kf}_i(C, d) \in \Gamma\}.$$

Then, extend this operator to $\mathcal{P}(\mathbf{Q})$ by $\mathrm{Cl}_i^\Gamma(C) := \bigcup\{\mathrm{Cl}_i^\Gamma(C_f) \mid C_f \subseteq_{fin} C\}$. When the context is clear, we may drop the superscript or subscript of Cl_i^Γ. Now this is a finitary closure operator as it satisfies, through the axioms of $\mathbb{LKVF}_m$,

$$\mathrm{Cl}(C) = \mathrm{Cl}(\mathrm{Cl}(C)), C \subseteq \mathrm{Cl}(C), C \subseteq D \Rightarrow \mathrm{Cl}(C) \subseteq \mathrm{Cl}(D).$$

When a set $C \subseteq \mathbf{Q}$ satisfies $C = \mathrm{Cl}(C)$, it is called a closed set. A classical result is that the collection of all closed sets under a closure operator forms a lattice $\langle L, \wedge, \vee\rangle$ with

$$\begin{aligned} L &= \{C \subseteq \mathbf{Q} \mid C = \mathrm{Cl}_i^\Gamma(C)\} \\ C \wedge D &= C \cap D \\ C \vee D &= \mathrm{Cl}_i^\Gamma(C \cup D), \end{aligned}$$

which we name $\mathfrak{L}_i^\Gamma$. For all $c \in \mathbf{Q}$, let $\mathrm{Cl}_i^\Gamma(c)$ stands for $\mathrm{Cl}_i^\Gamma(\{c\})$ to save a few brackets.

Also, given Γ, the indexed version of the propositional knowledge and the value knowledge of agent i is denoted by

$$K_{i,\Gamma} = \{\phi \mid K_i\phi \in \Gamma\}, \mathit{Kv}_{i,\Gamma} = \{d \mid \mathit{Kv}_i(d) \in \Gamma\}.$$

Then, it is not hard to see that $\mathfrak{L}_i^\Gamma$ is only dependent on $K_{i,\Gamma}$, i.e., if $K_{i,\Gamma} = K_{i,\Gamma'}$ then $\mathfrak{L}_i^\Gamma = \mathfrak{L}_i^{\Gamma'}$. This is because the closure operator Cl_i^Γ uses only the formulas of the form $\mathit{Kf}_i(C, d)$ in Γ, and if we assume $K_{i,\Gamma} = K_{i,\Gamma'}$,

$$\mathit{Kf}_i(C, d) \in \Gamma \Leftrightarrow K_i\mathit{Kf}_i(C, d) \in \Gamma \Leftrightarrow K_i\mathit{Kf}_i(C, d) \in \Gamma' \Leftrightarrow \mathit{Kf}_i(C, d) \in \Gamma'$$

for all $C \subseteq_{fin} \mathbf{Q}$ and $d \in \mathbf{Q}$.

For the completeness proof to go through, there is again a cardinality requirement for $\mathbf{G}$: $|\mathbf{G}| \geq |\mathbf{Q} \times \{0, 1\}|$, and without loss of generality, we identify $\mathbf{G}$ with $\mathbf{Q} \times \{0, 1\}$. The $\mathbf{Q}$ part will be used to construct the function domains and refute $\mathit{Kf}(C, d)$, while the $\{0, 1\}$ part will be used for refuting $\mathit{Kv}(d)$.

To use the $\mathbf{Q}$ part to construct the function domains, we need to forget the $\{0, 1\}$ part. Define function $\mathfrak{h}_i^\Gamma : \mathbf{Q} \times \{0, 1\} \to \mathfrak{L}_i^\Gamma$, $\langle c, n\rangle \mapsto \mathrm{Cl}_i^\Gamma(c)$ for each i, Γ. This map is forgetful about the second coordinate and turns a variable name into its closure. Again the superscript and subscript are dropped when no confusion arises. Now we are able to define a new version of the Lmax function set:

Definition 6 Given a maximal consistent set Γ and an agent index i, we can construct the dependency lattice $\mathfrak{L}$ and the corresponding $\mathfrak{h}$. Then define $F_i(\Gamma)$ to be the collection of all functions f on $\mathbf{G}$ with any arity $n \in \mathbb{N}$ such that:

$$\mathfrak{h}(f(x_1, x_2, \ldots x_n)) \leq \bigvee\{\mathfrak{h}(x_1), \mathfrak{h}(x_2), \ldots \mathfrak{h}(x_n)\},$$

where $\leq$ is defined in $\mathfrak{L}$ by $\mathfrak{a} \leq \mathfrak{b} \Leftrightarrow \mathfrak{a} \wedge \mathfrak{b} = \mathfrak{b}$, or equivalently, $\mathfrak{a} \subseteq \mathfrak{b}$. The empty disjunction is the bottom element of $\mathfrak{L}$: $\mathrm{Cl}(\emptyset)$.

It is straightforward to see that $F_i(\Gamma)$ is dependent only on $K_{i,\Gamma}$. Then we need to verify the soundness conditions immediately:

Proposition 8 For every maximal consistent set Γ and $i \in \mathbf{A}$, $F_i(\Gamma)$ contains all projection functions on $\mathbf{G}$ and is closed under composition.

Proof. Take a projection function $f(x_1, \ldots x_n) = x_k$. Then by the definition of join in a lattice,

$$\mathfrak{h}(x_k) \leq \bigvee\{\mathfrak{h}(x_1), \cdots, \mathfrak{h}(x_n)\}$$

since $\mathfrak{h}(x_k) \in \{\mathfrak{h}(x_1), \cdots, \mathfrak{h}(x_n)\}$.

For function composition, let $\overline{x}$ represent a sequence of variables and $\mathfrak{h}(\overline{x})$ the sequence after the application of $\mathfrak{h}$. Then take a function $f(\overline{x}) = g_0(g_1(\overline{x}_1), \cdots g_n(\overline{x}_n))$ where $\overline{x}$ includes the union of all $\overline{x}_k$s and all g functions are already in $F_i(\Gamma)$. Now

$$\begin{array}{rcl} \mathfrak{h}(f(\overline{x})) & = & \mathfrak{h}(g_0(g_1(\overline{x}_1), \cdots g_n(\overline{x}_n))) \\ & \leq & \bigvee\{\mathfrak{h}(g_1(\overline{x}_1), \cdots g_n(\overline{x}_n))\} \\ & \leq & \bigvee\{\vee\mathfrak{h}(\overline{x}_1), \cdots \vee \mathfrak{h}(\overline{x}_n)\} \\ & \leq & \bigvee \mathfrak{h}(\overline{x}). \end{array}$$

This shows that the composition f satisfies the requirement and is in $F_i(\Gamma)$. ■

The next proposition shows why we use the dependence lattice to define the function domains for each agent. The proposition says that to make $Kf_i(C, d)$ true, we only need to make sure that functionality holds, and to make $Kf_i(C, d)$ false, we do not need to pay any special attention as the function domain $F_i(\Gamma)$ has already taken care of everything.

Proposition 9 For every $\sigma \in 2^{\mathbf{Q}}$, define $v_\sigma : \mathbf{Q} \to \mathbf{G}, d \mapsto \langle d, \sigma(d)\rangle$. This means we restrict the value of $d \in \mathbf{Q}$ to be $\langle d, 0\rangle$ or $\langle d, 1\rangle$. Now for every maximal consistent set Γ, $i \in \mathbf{A}$, $C \subseteq_{fin} \mathbf{Q}$, $d \in \mathbf{Q}$, and $\Sigma \subseteq 2^{\mathbf{Q}}$:

- if Σ satisfies the functionality condition for C, d, namely for all $\sigma_1, \sigma_2 \in \Sigma$, $\sigma_1(C) = \sigma_2(C)$ implies $\sigma_1(d) = \sigma_2(d)$, and if $Kf_i(C, d) \in \Gamma$, then there exists $f \in F_i(\Gamma)$ such that for all $\sigma \in \Sigma$, $v_\sigma(d) = f(v_\sigma(C))$;
- if $Kf_i(C, d) \notin \Gamma$ then for all $\sigma \in \Sigma$ and for all $f \in F_i(\Gamma)$, $v_\sigma(d) \neq f(v_\sigma(C))$.

Proof. First notice that in the definition of $F_i(\Gamma)$, the restriction actually forgets the second coordinate of the inputs and outputs. But it is the second coordinate that all $\sigma \in \Sigma$ try to adjust. By definition, the first coordinates of $v_\sigma(c)$ for all $c \in \mathbf{Q}$ are just themselves. So for all $c \in \mathbf{Q}, \sigma \in \Sigma$, $\mathfrak{h}(v_\sigma(c)) = \mathrm{Cl}(c)$.

If $Kf_i(C, d) \in \Gamma$, then (dropping the super and subscripts) $d \in \mathrm{Cl}(C)$. This means the same as $\{d\} \subseteq \mathrm{Cl}(C)$, which, by the fact that Cl is a closure operator, implies $\mathrm{Cl}(d) \subseteq \mathrm{Cl}(\mathrm{Cl}(C)) = \mathrm{Cl}(C)$. Then $\mathrm{Cl}(d) \subseteq \mathrm{Cl}(C)$, which means $\mathfrak{h}(v_\sigma(d)) \leq \mathrm{Cl}(C)$ in $\mathfrak{L}$ for all $\sigma \in \Sigma$. Also, $\mathrm{Cl}(C) = \bigvee\{\mathrm{Cl}(c_1), \mathrm{Cl}(c_2), \dots \mathrm{Cl}(c_n)\} = \bigvee \mathfrak{h}(v_\sigma(C))$ for all $\sigma \in \Sigma$. So indeed $\mathfrak{h}(v_\sigma(d)) \leq \bigvee \mathfrak{h}(v_\sigma(C))$ in $\mathfrak{L}$. Together with the functionality assumed for Σ, this means mapping $v_\sigma(C)$ to $v_\sigma(d)$ simultaneously for all $\sigma \in \Sigma$ is allowed in $F_i(\Gamma)$. Then we can extend this partial map to a map from $\mathbf{G}^n$ to $\mathbf{G}$ in $F_i(\Gamma)$. An easy solution is to do projection for all other possible inputs.

If $Kf_i(C, d) \notin \Gamma$, then $d \notin \mathrm{Cl}(C)$ and hence $\mathrm{Cl}(d) \nsubseteq \mathrm{Cl}(C)$. If Σ is empty, the statement is trivially true. So assume Σ is not empty. Now take an arbitrary $\sigma \in \Sigma$. Then $\mathfrak{h}(v_\sigma(d)) \nleq \bigvee \mathfrak{h}(v_\sigma(C))$, which violates the restriction on $F_i(\Gamma)$ if $v_\sigma(C)$ is to be mapped to $v_\sigma(d)$. Thus, for all $f \in F_i(\Gamma), v_\sigma(d) \neq f(v_\sigma(C))$.

This proposition says that the dependency lattice $\mathfrak{L}_i^\Gamma$ and the corresponding function domain $F_i(\Gamma)$ form a suitable representation of the function domain that i uses implicitly given i's knowledge and ignorance in Γ. As we hinted before the construction, this function domain is so specific about what is possible that when $Kf_i(C, d)$ is not known, it is not rejected by a failure of functionality, which requires at least two epistemically possible assignment, but by a failure of conforming to the prior knowledge encoded in the function domain, as shown by the second bullet in the previous proposition. On the other hand, once functionality holds in all possible assignments, we do not need to worry about whether the function domain allows it or not, which is clear from the proof of the first bullet. Thus, this $F_i(\Gamma)$ is a perfect choice.

For the Kv_i part, we need to adjust the assignments of variables to construct more (epistemically) possible assignments to reject formulas like $Kv_i(d)$ which is not in Γ: if in one world d is assigned to be x, then we want to make an adjustment to get a new world where it is assigned to $y \neq x$. This will be done by moving the value of d to $\langle d, 1\rangle$ from $\langle d, 0\rangle$ or vice versa. And for agent i in a maximal consistent set Γ, the variables to be moved are exactly $\overline{Kv_{i,\Gamma}} = \{d \mid Kv_i(d) \notin \Gamma\}$, the complement of the set of the variables with a known value by i. By maximality, it is also the collection of all $d \in \mathbf{Q}$ such that $\neg Kv_i(d) \in \Gamma$. It is crucial to move the value of all variables in $\overline{Kv_{i,\Gamma}}$ at once, as otherwise there might be some unwanted violation of functionality: even though for both $\sigma = \sigma_1, \sigma_2$, $\mathfrak{h}(v_\sigma(d)) \leq \bigvee \mathfrak{h}(v_\sigma(C))$, it could be that $v_{\sigma_1}(C) = v_{\sigma_2}(C)$ while $v_{\sigma_1}(d) \neq v_{\sigma_1}(d)$. So in this case, no functional dependency exists from C to d, but the reason is not that d is at the wrong place in the lattice, but instead the failure of functionality. We must avoid this situation, by changing all values of variables in $\overline{Kv_{i,\Gamma}}$ simultaneously when producing a new possible assignments in a new possible world. This motivates the following definition:

Definition 7 (Value Move) Given Γ a maximal consistent set and $i \in \mathbf{A}$, define the value move operator $\mathrm{Mv}_i^\Gamma : 2^{\mathbf{Q}} \to 2^{\mathbf{Q}}$:

$$\mathrm{Mv}_i^\Gamma(\sigma)(d) = \begin{cases} \sigma(d) & d \in Kv_{i,\Gamma} \\ 1 - \sigma(d) & d \in \overline{Kv_{i,\Gamma}}. \end{cases}$$

This operator captures agent i's switching of the values of the variables in $\overline{Kv_{i,\Gamma}}$ all at once. Two important properties should be noted. First, Mv_i^Γ is dependent only on $K_{i,\Gamma}$. Indeed it only depends on $Kv_{i,\Gamma}$ but because of the axioms `KV4` and `KV5`, it is equivalent to say that it depends only on $K_{i,\Gamma}$. This means that if $K_{i,\Gamma} = K_{i,\Gamma'}$, then as an operator, $\mathrm{Mv}_i^\Gamma = \mathrm{Mv}_i^{\Gamma'}$.

Another important property of this operator is that $\mathrm{Mv}_i^\Gamma(\mathrm{Mv}_i^\Gamma(\sigma)) = \sigma$ for all Γ, i, σ ranging over maximal consistent sets, $\mathbf{A}$ and $2^{\mathbf{Q}}$. Thus, it is actually an inverse operator.

Equipped with the above definitions, the canonical model can now be defined:

Definition 8 (Canonical Model) Build a model $\mathcal{M} = \langle W, \langle \sim_i \rangle_{i \in \mathbf{A}}, U, V, \langle \mathbf{F}_i \rangle_{i \in \mathbf{A}} \rangle$ as follows:

- $W = \{\langle \Gamma, \sigma \rangle \mid \Gamma \text{ a maximal consistent set}, \sigma \in 2^{\mathbf{Q}}\}$,
- $\langle \Gamma, \sigma \rangle \sim_i \langle \Gamma', \sigma' \rangle$ iff
  1. $K_{i,\Gamma} = K_{i,\Gamma'}$, which says that two worlds must share the same set of knowledge of i, and
  2. $\sigma = \sigma'$ or $\sigma = \mathrm{Mv}_i^\Gamma(\sigma')$, which says that any agent i needs to see some different possible assignments of the variables, but not too many: just two,
- $U(\langle \Gamma, \sigma \rangle, p) = [p \in \Gamma]$,
- $V(\langle \Gamma, \sigma \rangle, d) = \langle d, \sigma(d) \rangle$, or equivalently using notations introduced above in Proposition 9, $V(\langle \Gamma, \sigma \rangle) = v_\sigma$,
- $\mathbf{F}_i(\langle \Gamma, \sigma \rangle) = F_i(\Gamma)$.

Before proving the truth lemma, it must be shown that $\mathcal{M}$ is indeed a model of $\mathbf{LKVF}_m^*$. This amounts to checking the following:

- $\sim_i$ is an equivalence relation for all $i \in \mathbf{A}$,
- $\mathbf{F}_i(\langle \Gamma, \sigma \rangle)$ satisfies the soundness condition,
- if $\langle \Gamma, \sigma \rangle \sim_i \langle \Gamma', \sigma' \rangle$ then $\mathbf{F}_i(\langle \Gamma, \sigma \rangle) = \mathbf{F}_i(\langle \Gamma', \sigma' \rangle)$.

Because $\sim_i$ is defined using equality, its reflexivity is easy to see. We need the two special properties of Mv_i^Γ noted right after the Definition 7 to show symmetry and transitivity.

For symmetry, suppose $\langle \Gamma, \sigma \rangle \sim_i \langle \Gamma', \sigma' \rangle$. Then $K_{i,\Gamma} = K_{i,\Gamma'}$. Thus, $\mathrm{Mv}_i^\Gamma = \mathrm{Mv}_i^{\Gamma'}$ and $\sigma = \mathrm{Mv}_i^\Gamma(\sigma') = \mathrm{Mv}_i^{\Gamma'}(\sigma')$. Also, as $\mathrm{Mv}_i^{\Gamma'}$ is an inverse operator, by applying it twice, we get $\sigma' = \mathrm{Mv}_i^{\Gamma'}(\sigma)$. So it can be concluded that $\langle \Gamma', \sigma' \rangle \sim_i \langle \Gamma, \sigma \rangle$.

Transitivity can be shown similarly. Suppose $\langle \Gamma_1, \sigma_1 \rangle \sim_i \langle \Gamma, \sigma \rangle \sim_i \langle \Gamma_2, \sigma_2 \rangle$. It immediately follows that $\mathrm{Mv}_i^{\Gamma_1} = \mathrm{Mv}_i^\Gamma = \mathrm{Mv}_i^{\Gamma_2}$. So we can treat all of them as Mv_i^Γ. Then we know $\sigma = \sigma_1$ or $\sigma = \mathrm{Mv}_i^\Gamma(\sigma_1)$, and $\sigma = \sigma_2$ or $\sigma = \mathrm{Mv}_i^\Gamma(\sigma_2)$. There are in total four possibilities depending on which disjuncts hold, and the only less trivial one is when $\sigma_1 = \mathrm{Mv}_i^\Gamma(\sigma)$ and $\sigma = \mathrm{Mv}_i^\Gamma(\sigma_2)$. But if that is the case, then $\sigma_1 = \mathrm{Mv}_i^\Gamma(\mathrm{Mv}_i^\Gamma(\sigma_2)) = \sigma_2$. So transitivity holds.

The soundness condition was already shown when $\mathbf{F}_i(\langle \Gamma, \sigma \rangle) = F_i(\Gamma)$ is defined in Proposition 8. We also noted that $F_i(\Gamma)$ only depends on $K_{i,\Gamma}$ because it only depends on the dependency lattice $\mathfrak{L}_i^\Gamma$, which in turn only depends on $K_{i,\Gamma}$. If $\langle \Gamma, \sigma \rangle \sim_i \langle \Gamma', \sigma' \rangle$, $K_{i,\Gamma} = K_{i,\Gamma'}$ and $F_i(\Gamma) = F_i(\Gamma')$, so indeed $\mathbf{F}_i(\langle \Gamma, \sigma \rangle) = \mathbf{F}_i(\langle \Gamma', \sigma' \rangle)$. So we conclude that $\mathcal{M}$ is a model of $\mathbf{LKVF}_m^*$.

The unconventional second condition for $\sim_i$ is there for the purpose of preventing unwanted failure of functionality. As explained after Proposition 9, we are not refuting $Kf(C, d)$ using functionality, so it is better to keep the functionalities between as many variables as possible. In particular, all functionalities between the variables in $\overline{Kv_{i,\Gamma}}$ can be preserved. The condition does this by requiring that if i sees more than one possibility for some variables, then all the values of $\overline{Kv_{i,\Gamma}}$ must change to a different epistemic possibility together using the value move operator. This makes impossible the situation where one variable in $\overline{Kv_{i,\Gamma}}$ realizes a different possibility while another stays the same, a situation that characterizes the failure of functionality.

Now the truth lemma in this case can be proven:

Lemma 4 (Truth Lemma) For all ϕ in the language of $\mathbf{LKVF}^*_m$ and all maximal consistent sets Γ in the axiom system $\mathbb{LKVF}_m$, $\mathcal{M}, \langle\Gamma, \sigma\rangle \vDash \phi$ if and only if $\phi \in \Gamma$.

Proof. Use induction on ϕ. The propositional letters and boolean combination cases are conventional. We focus on the knowledge cases.

$\phi = K_i\psi$. If $K_i\psi \in \Gamma$, then by the definition of $\sim_i$, for all $\langle\Gamma', \sigma'\rangle \sim_i \langle\Gamma, \sigma\rangle$, $K_{i,\Gamma} = K_{i,\Gamma'}$. Thus, $\psi \in K_{i,\Gamma'}$ and $K_i\psi \in \Gamma'$. By axiom $\mathtt{T}$, $\psi \in \Gamma'$, and using the induction hypothesis, $\mathcal{M}, \langle\Gamma', \sigma'\rangle \vDash \psi$. Thus, $\mathcal{M}, \langle\Gamma, \sigma\rangle \vDash K_i\psi$ by the semantic clause of K_i.

If $K_i\psi \notin \Gamma$, then by a standard argument using axioms and the maximality of Γ, $K_{i,\Gamma} \cup \{\neg\psi\}$ is consistent and expandable to a maximal consistent set Γ'. Then $\langle\Gamma', \sigma\rangle \sim_i \langle\Gamma, \sigma\rangle$ and $\mathcal{M}, \langle\Gamma, \sigma\rangle \vDash \neg\psi$ by the induction hypothesis. So $\mathcal{M}, \langle\Gamma, \sigma\rangle \nvDash K_i\psi$.

$\phi = Kv_i(d)$. If $Kv_i(d) \in \Gamma$, then $d \in Kv_{i,\Gamma}$ and thus $\mathrm{Mv}_i^\Gamma(\sigma)(d) = \sigma(d)$. Now for all $\langle\Gamma', \sigma'\rangle \sim_i \langle\Gamma, \sigma\rangle$, σ is equal to σ' or $\mathrm{Mv}_i^\Gamma(\sigma')$. But as $d \in Kv_{i,\Gamma}$, Mv_i^Γ is not changing the value of d. So in either case, $\sigma'(d) = \sigma(d)$. Thus, the value of d is fixed to $\langle d, \sigma(d)\rangle$ among all worlds accessible by i from $\langle\Gamma, \sigma\rangle$.

If $Kv_i(d) \notin \Gamma$, then $d \notin Kv_{i,\Gamma}$ and $\mathrm{Mv}_i^\Gamma(\sigma)$ will change the value of d. Take the world $\langle\Gamma', \sigma'\rangle$ with $\sigma' = \mathrm{Mv}_i^\Gamma(\sigma)$. Then $\sigma = \mathrm{Mv}_i^\Gamma(\sigma')$, so $\langle\Gamma, \sigma\rangle \sim_i \langle\Gamma', \sigma'\rangle$. Also, $\sigma'(d) = 1 - \sigma(d) \neq \sigma(d)$. Thus, $V(\langle\Gamma', \sigma'\rangle, d) \neq V(\langle\Gamma, \sigma\rangle, d)$. By the semantic clause of $Kv_i(d)$, $\mathcal{M}, \langle\Gamma, \sigma\rangle \nvDash Kv_i(d)$.

$\phi = Kf_i(C, d)$. Suppose $Kf_i(C, d) \in \Gamma$. Then we should first show that the functionality condition holds. For any $\langle\Gamma_1, \sigma_1\rangle, \langle\Gamma_2, \sigma_2\rangle \sim_i \langle\Gamma, \sigma\rangle$, if $V(\langle\Gamma_1, \sigma_1\rangle, C) = V(\langle\Gamma_2, \sigma_2\rangle, C)$, then there are two possibilities

- $C \subseteq Kv_{i,\Gamma}$. Then by axiom $\mathtt{VF}$, $d \in Kv_{i,\Gamma}$ as well, and by the argument in the previous case, $V(\langle\Gamma_1, \sigma_1\rangle, d) = V(\langle\Gamma_2, \sigma_2\rangle, d) = \langle d, \sigma(d)\rangle$.
- $C \nsubseteq Kv_{i,\Gamma}$. Then take $c \in C \cap \overline{Kv_{i,\Gamma}}$. Since $V(\langle\Gamma_1, \sigma_1\rangle, C) = V(\langle\Gamma_2, \sigma_2\rangle, C)$, $\sigma_1(c) = \sigma_2(c)$. Because $\langle\Gamma_1, \sigma_1\rangle \sim_i \langle\Gamma, \sigma\rangle \sim_i \langle\Gamma_2, \sigma_2\rangle$, $\langle\Gamma_1, \sigma_1\rangle \sim_i \langle\Gamma_2, \sigma_2\rangle$. So either $\sigma_1 = \sigma_2$ or $\sigma_1 = \mathrm{Mv}_i^{\Gamma_1}(\sigma_2)$. But the latter case cannot happen because if that is true, then $\sigma_1(c) \neq \sigma_2(c)$ since $c \in \overline{Kv_{i,\Gamma}}$. So $\sigma_1 = \sigma_2$ and in particular $\sigma_1(d) = \sigma_2(d)$. Thus, $V(\langle\Gamma_1, \sigma_1\rangle, d) = V(\langle\Gamma_2, \sigma_2\rangle, d)$.

Indeed, by our definition of $\sim_i$, among all worlds accessible from $\langle\Gamma, \sigma\rangle$ by $\sim_i$, there are altogether only two possible valuations: σ and $\mathrm{Mv}_i^\Gamma(\sigma)$. Thus, by applying Proposition 9 to set $\Sigma = \{\sigma' \mid \langle\Gamma', \sigma'\rangle \sim_i$

$\langle \Gamma, \sigma \rangle\}$, it follows that there exists a function $f \in F_i(\Gamma) = \mathbf{F}_i(\langle \Gamma, \sigma \rangle)$ such that $V(\langle \Gamma', \sigma' \rangle, d) = f(V(\langle \Gamma', \sigma' \rangle, C))$ for all $\langle \Gamma', \sigma' \rangle \sim_i \langle \Gamma, \sigma \rangle$. So $\mathcal{M}, \langle \Gamma, \sigma \rangle \vDash Kf_i(C, d)$.

If $Kf_i(C, d) \notin \Gamma$, then by Proposition 9 again, for every function $f \in F_i(\Gamma) = \mathbf{F}_i(\langle \Gamma, \sigma \rangle)$, there exists $\langle \Gamma', \sigma' \rangle \sim_i \langle \Gamma, \sigma \rangle$ such that $V(\langle \Gamma', \sigma' \rangle, d) \neq f(V(\langle \Gamma', \sigma' \rangle, C))$. Actually $\langle \Gamma, \sigma \rangle$ itself works here. Thus, $\mathcal{M}, \langle \Gamma, \sigma \rangle \nvDash Kf_i(C, d)$. ■

From the truth lemma, it can be concluded that every consistent set is satisfied somewhere in the canonical model $\mathcal{M}$ built above. So the completeness of $\mathbb{LKVF}$ with respect to $\mathbf{LKVF}^*_m$ follows. Together with the soundness proven in Proposition 8, we obtain an axiomatization of $\mathbf{LKVF}^*_m$:

Theorem 5 Under the cardinality requirement $\mathbf{G} \geq |\mathbf{Q} \times \{0, 1\}|$, $\mathbb{LKVF}_m$ is an axiomatization of $\mathbf{LKVF}^*_m$.

7 Discussion and Future Work

First, we discuss the semantics of the *Kf* operator. Obviously, while $Kv(d)$ means that there is only one value for d to take, in general, the truth of $Kf(C, d)$ does not force the set of possible functional dependency relations of d on C to be a singleton.

It could be argued that the agent can nevertheless regard all those candidates as equivalent, because they must have exactly the same behavior over the partial domain $P = \{V(w, C) \mid w \in W\}$. And things in $\mathbf{G}^{|C|}$ but outside this set P are epistemically impossible. Thus, the behavior of functions on $\mathbf{G}^{|C|} \backslash P$ is something that our agent can and will ignore if situations epistemically impossible do not concern the agent. One example, also mentioned in the introduction, is when "knowing-value" is the real objective of the agent and "knowing-dependency" only expresses the agent's potential to know more values. The semantics proposed in this paper allows adjustments to $\mathbf{F}$, which might be a consequence of an agent's concern about situations epistemically impossible, but not necessarily. And even if it is the case, the semantics does not show how $\mathbf{F}$ is derived from what concerns of the agents.

It is not uncommon that epistemic possibilities are not the right place to stop when evaluating knowledge of functional dependency. Consider the following example:

> I know the color of my hair. Therefore, I know the color of my hair functionally depends on the number of fingers I have.

This argument is very hard to swallow intuitively. Yet it is validated by the axiom `EXT`. Indeed, in the current setting of the semantics of *Kf*, to validate this, we only need to allow a moderate amount of constant functions in our function domain. The root of the problem is that, in a pure epistemic logic setting, if something is known, the agent has no access to other alternatives as knowledge is the only modality here, whereas in most realistic situations, even when something is known, we have modal access to some possibilities different from the known one. For example, possibilities in the future or past can be used to explain why the color of my hair is not really dependent on the number of fingers I have. And even when I have not and will not change the color of my hair, we can still use metaphysical possibilities: "the color of my hair *could* be different, regardless of how many fingers I have."

Thus, it might be of interest to capture knowledge of functional dependency in another modality. To do this we can add a new modality $\Box$ interpreted by a relation R. Then "knowing a/the functional dependency" can now be expressed by an operator $Kf^{\Box}$ with the following semantics:

$$\langle W, \sim, R, U, V \rangle, w \vDash Kf^{\Box}(C, d) \Leftrightarrow \exists f \in \mathbf{F}, \forall w' \sim w, \forall w^*, Rw'w^* \Rightarrow V(w^*, d) = f(V(w^*, C))$$

where $\sim$ is the epistemic indistinguishability relation. This definition still says that there exists a function that works for all epistemically indistinguishable worlds. But here "works" means f captures the functional dependency of d upon C with respect to another modality $\Box$ which might be different from K.

The choice of R can be arbitrary, but at least two interesting candidates are immediate: an equivalence relation to capture metaphysical possibilities and a linear or branching time relation used in temporal logics. A simple observation is that, if we still want a new version of VF, namely

$$\mathtt{VF}' : \bigwedge_{c \in C} Kv(c) \wedge Kf^{\Box}(C, d) \to Kv(d)$$

to be valid, we need R to be reflexive. Otherwise, the functional dependency might be only talking about worlds far away from the actual world, though accessible through R. Since the choice for R can be flexible, there will be many interesting results to be discovered under this semantics. In particular, for the study of completeness, we might want to add more first order features to facilitate a proof more similar to its first order counterpart, a strategy successfully employed in [1]. It might be desirable because, with two modalities, a direct construction of value assignments can be unmanageable.

But a demanding reader may still not be satisfied, as even if we add a new modality, the choice of the functions could be nonunique again. This motivates another interpretation of knowledge of functional dependency, emphasizing even more the "knowledge" part: $Kf(C, d)$ says that the agent has gathered so much information that there is (almost) exactly one function that can be used to explain the data he/she has seen so far. Thus, knowledge appears only when there is only one possible or a few very plausible explanations. If there is no possible explanation in the sense that no function in the function domain $\mathbf{F}$ is applicable, or there are too many explanations, no knowledge is obtained. This sounds natural, but much more technically will be needed to formalize this: either a counting operator, or a probabilistic operator tracking the posterior distribution over the candidate explanations.

There are also interesting possible extensions of the framework given in this paper. For example, the multiagent case here assumed a no-interaction semantics. But once we require prior knowledge of possible functions to be available to other agents, interesting interactions will appear. For example, suppose $\mathbf{F}_j$ is known to agent i, i.e., if $w \sim_i w'$ then $\mathbf{F}_j(w) = \mathbf{F}_j(w')$. Then the following is valid:

$$Kv_i(c) \wedge Kv_i(d) \wedge K_i(Kv_j(c) \wedge Kv_j(d)) \to K_i Kf_j(c, d) \vee K_i \neg Kf_j(c, d).$$

Intuitively this says that if agent i knows the values of c, d and knows that agent j knows, then either i knows that j has an explanation of the value of c, d or i knows that j does not have one. The antecedent fixes the value of c, d in all worlds accessible first from i and then from j. Thus if j fails or succeeds to explain this particular instance, agent i knows it. Stronger interactions will appear if we require all agents to share a single prior knowledge base $\mathbf{F}$, i.e., for all i, w, $\mathbf{F}_i(w) = \mathbf{F}$. Then the following is valid:

$$Kv_i(c) \wedge Kv_i(d) \wedge K_i(Kv_j(c) \wedge Kv_j(d)) \to (Kf_i(c,d) \to Kf_j(c,d)).$$

This says that if i knows the value of c, d and knows that j knows them, then i being able to explain this instance implies that j can explain it as well. To axiomatize these two cases, new axioms and techniques will emerge. Further, we can also add an operator that expresses knowledge about other agents' function domain.

Computationally, we see without too much surprise that the finite model property holds. For all the three single agent cases with a finite language, the required size of $\mathbf{G}$ and the size of the model constructed can be explicitly computed. In the multiagent case, a standard filtration method can also be applied quite straightforwardly. Notice that in each of the three cases, the completeness proof requires a minimal size of $\mathbf{G}$. A natural question is whether we can bring down the size requirement by giving more economic completeness proofs. In particular, the double exponential size requirement in the single agent fixed intermediate function domain case seems to be too large, while the number of value assignments seems too small (just 2). We might be able to implement a trade-off here or a smarter lattice construction.

In summary, introducing knowledge about functional dependency relations brings us ample new opportunities to extend the border of epistemic logic. There will be a lot more to achieve.

References

[1] Alexandru Baltag. To know is to know the value of a variable. In *Advances in Modal Logic 11*, pages 135–155, 2016.

[2] Tao Gu and Yanjing Wang. "knowing value" logic as a normal modal logic. In *Advances in Modal Logic 11*, pages 362–381, 2016.

[3] Jaakko Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca N.Y., 1962.

[4] J. A. Plaza. Logics of public communications. In M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, and Z. W. Ras, editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216, 1989.

[5] Jouko Väänänen et al. Modal dependence logic. *New perspectives on games and interaction*, 4:237–254, 2008.

[6] Jouko A. Väänänen. *Dependence Logic - A New Approach to Independence Friendly Logic*, volume 70 of *London Mathematical Society student texts*. Cambridge University Press, 2007.

[7] Jan van Ejick, Malvin Gattinger, and Yanjing Wang. Knowing values and public inspection. In *to appear in Proceedings of ICLA2017*, 2017.

[8] Yanjing Wang. Beyond knowing that: a new generation of epistemic logics. In Gabriel Sandu Hans van Ditmarsch, editor, *Jaakko Hintikka on knowledge and game theoretical semantics*, page to appear. Springer, 2016.

[9] Yanjing Wang and Jie Fan. Knowing that, knowing what, and public communication: Public announcement logic with Kv operators. In *Proceedings of International Joint Conferences on Artificial Intelligence*, pages 1139–1146, 2013.

[10] Yanjing Wang and Jie Fan. Conditionally knowing what. In *Advances in Modal Logic 10, invited and contributed papers from the tenth conference on "Advances in Modal Logic," held in Groningen, The Netherlands, August 5-8, 2014*, pages 569–587, 2014.

[11] Armstrong W. William. Dependency structures of data base relationships. In *Information Processing 74, Proceedings of IFIP Congress 74*, pages 580–583, 1974.