Knowing Correlations
(Extended Abstract)

ALEXANDRU BALTAG
ILLC, University of Amsterdam

Informationally, a question can be encoded as a variable, taking various values ("answers") in different possible worlds. If, in accordance to the recent trend towards an interrogative epistemology, "To know is to know the answer to a question" [6], then we are lead to paraphrasing Quine's motto: *To know is to know the value of a variable* [5].

In a paper with this title [2], presented at "Advances in Modal Logic" 2016, I introduced a logical account of this form of knowledge, by building on the work of Plaza [4], and Wang and Fan [7] on capturing 'knowledge de re' (knowledge of an object, "knowledge what") over Kripke models. In its turn, this research is motivated by work in Security on knowledge of keys and passwords. A *possible world* in these models is the same as a first-order assignment, i.e. an assignment $w : Var \to D$ of values (from some fixed domain D) to variables $x \in Var$. Plaza's semantics for 'knowledge de re' is given by putting: $w \models K_a x$ iff $\forall v \sim_a w \ (v(x)) = w(x))$. This is a natural analogue of the usual semantics of "knowledge that" in epistemic logic: an agent knows the value of x if *that value is the same in all her epistemic alternatives.* This is the sense in which may say that (s)he knows a password or an encryption key.

However, questions are never investigated in isolation: we answer questions by reducing them to other questions. This means that the proper object of knowledge is uncovering *correlations* between questions. *To know is to know a functional dependence between variables.* In the same paper, I investigated (and completely axiomatized) a Logic of Epistemic Dependency, that can express knowledge of functional dependencies between (the values of) variables, as well as dynamic modalities for learning new such dependencies. This dynamics captures the widespread view of knowledge acquisition as a process of learning correlations (with the goal of eventually tracking causal relationships in the actual world). This research is also clearly related to the study of dependencies in Database theory: indeed, the well-known Armstrong axioms [1] for dependency are provable in the Logic of Epistemic Dependency.

For a string $\mathbf{y} = (y_1, \ldots, y_n)$ of variables and a variable x, the *epistemic* *dependency formula* $K_a^{\mathbf{y}}x$ says that an agent knows the value of X *conditional* on being given the values of variables $\mathbf{y}$. The semantics is the obvious generalization of the above clause: if we use the abbreviation $w(\mathbf{y}) = v(\mathbf{y})$ for the conjunction $w(y_1) = v(y_1) \wedge \ldots w(y_n) = v(y_n)$, then we put

$$w \models K_a^{\mathbf{y}}x \quad \text{iff} \quad \forall v \sim_a w \; (w(\mathbf{y}) = v(\mathbf{y}) \Rightarrow v(x) = w(x)) \, .$$

In words: an agent knows x given $\mathbf{y}$ if the value of x is the same in all the epistemic alternatives that agree with the actual world on the values of $\mathbf{y}$.

In this paper I propose an extension of the Epistemic Dependence Logic, called the Logic of Correlations (LC). Essentially, it uses van Benthem's Generalized-Assignment Semantics of First-Order Logic (on *dependency models*, also known as general assignment models). While the usual semantics of FOL allows all possible assignments of variables to objects, a dependency model restricts the range of available assignments to a given subset. Dependency models were proposed by van Benthem as a model for capturing dependencies and correlations between variables. It is known that FOL with this generalized semantics becomes decidable [3]. Unfortunately, the resulting logic is so weak that it can no longer express variable dependencies in an explicit manner (though the model still implicitly captures them).

Our Logic of Correlations proposes a remedy to this limitation, by adding to the syntax functional terms of the form $x(\mathbf{y}, \phi)$, denoting *the unique (value of) x determined by (the values of) $\mathbf{y}$ (in the current assignment) and by condition ϕ* (if such an x exists). For a given variable assignment w, this term has a denotation only if it is indeed the case that the current values $w(\mathbf{y})$ of variables $\mathbf{y}$, together with the information that ϕ, uniquely determine a specific value of x. So the precondition for denoting is a statement $x(\mathbf{y}, \phi) \downarrow$ (definable in our logic), saying that $\mathbf{y}$ *uniquely determines x conditional on ϕ*. When thinking of y as an agent (or group of agents), we may alternatively write this precondition as an "epistemic" statement $K_{\mathbf{y}}^{\phi}x$, saying that *agent (group) $\mathbf{y}$ 'knows' (the value of) x conditional on ϕ*. (This was the interpretation I adopted when introducing Epistemic Dependence Logic in my AiML paper.) By taking ϕ to be any tautological formula $\top$, we obtain a formula $K_{\mathbf{y}}x$ that captures a 'local' version of the so-called dependence atoms from Dependence Logic: all assignments (available in the model) that agree with the current assignment on $\mathbf{y}$ also agree with it on x. The usual "global" dependency statement $=(\mathbf{y}, x)$ from Dependence Logic (saying that every two available assignments that agree on $\mathbf{y}$ also agree on x) can be defined using the local version (as $KK_{\mathbf{y}}x$, where the first K is just the global modality, i.e. the universal quantifier over all the available variables).

Our first important result in this paper is that the Logic of Correlations is decidable. We show this by using the quasi-model method pioneered by van Benthem (and closely related to the mosaic method of Andreka and

Nemeti). We also look at its embedding into fragments of FOL with the standard semantics. It is well known that FOL with generalized assignment semantics can be embedded in the Guarded Fragment (of FOL) with standard semantics, and moreover that the same quasi-model techniques can be used to show the decidability of the Guarded Fragment (Andreka et alia), and of further extensions (the so-called Packed Fragment). So it is natural to ask whether a similar embedding can be found for our Logic of Correlations.

The second main contribution in this paper is to embedd LC into a decidable extension of the Guarded Fragment (on standard first-order models), which we call the Guarded Definite-Description Fragment. This is obtained by adding *guarded definite descriptions* of the form $!x_i.\exists \mathbf{y}(P\mathbf{xy} \wedge \phi\mathbf{xy})$ (where x_i is one of the x's), denoting *the unique* x_i *satisfying the guardedly-quantified formula* $\exists \mathbf{y}(P\mathbf{xy} \wedge \phi\mathbf{xy})$ (if such an x exists). Our quasi-model technique can be easily adapted to prove decidability of this fragment.

A natural open question is whether these decidability results can be further extended, say by adding "packed definite descriptions" to the Packed Fragment, and use the mozaic method. We conjecture that the answer is yes. Further fixed-point extensions are also worth considering. It is known that the extensions of Guarded and Packed fragments with monotonic fixed points are still decidable. What about the fixed-point extensions of the logics considered in this paper?

References

[1] W. J. Armstrong, Dependency structures of data base relationships, IFIP Congress, vol. 74, pp. 580-583, 1974.

[2] A. Baltag, To Know is to Know the Value of a Variable. In Proceedings of AiML '16 (Budapest, 2016), *Advances in Modal Logic* vol. 11, pp. 135-155, College Publications 2016.

[3] H. Andreka, J. van Benthem, and I. Nemeti. Modal languages and bounded fragments of predicate logic. Journal of Philosophical Logic, vol. 27, nr. 3, pp 217274, 1998.

[4] J. Plaza, Logics of public communication, Synthese, vol. 158, pp. 165-179, 2007.

[5] W.V. Quine, On What There Is, Review of Metaphysics, vol. 3, nr. 5, pp 21-36, 1948.

[6] J. Schaffer, Knowing the answer, Philosophy and Phenomenological Research, vol. 75, nr. 2, pp 383-403, Wiley Online Library 2007.

[7] Wang, Y. and Fan, J., Conditionally knowing what, Advances in Modal Logic'14, pp. 569-587, College Publications, 2014.