Matthew effects via dependence and independence logic (work in progress)

Sabine Frittella[1], Giuseppe Greco[2], Michele Piazzai[2], Nachoem Wijnberg[3,4], and Fan Yang[2]

[1] Laboratoire d'Informatique Fondamentale d'Orléans, France
[2] Delft University of Technology, The Netherlands
[3] University of Amsterdam, The Netherlands
[4] University of Johannesburg, South Africa

Matthew effect. Introduced by Merton [5], the term *Matthew effect* was used in reference to the self-reinforcing process whereby reputationally rich academics tend to get richer over time. The author defined this phenomenon as 'the accruing of large increments of peer recognition to scientists of great repute for particular contributions in contrast to the minimizing or withholding of such recognition for scientists who have not yet made their mark'. Its recurring appearance in social life led to its recognition as a powerful engine of social, economic, and cultural inequality, to the extent that it can be considered a social law.

There is an extensive literature on the Matthew effect in the fields of sociology, economics, and management [1]. Yet, the effect is not precisely and unequivocally defined in the literature: as a result, researchers are hardly able to compare or integrate theoretical models and empirical findings. This motivates our present attempt to formalize the Matthew effect through mathematical logic.

Dependence and independence logic. The logical framework that we propose for the formalization is the framework of *Dependence and independence logic* introduced by Väänänen [6] and by Grädel and Väänänen [2]. This framework aims at characterizing the notions of *dependence* and *independence* found in social and natural sciences, such as the dependencies involved in Matthew effects. The logics extend first-order logic with new atomic formulas, called *dependence and independence atoms*, that specify explicitly the dependence and independence relations between variables. To evaluate formulas concerning dependency statements the logics adopt an innovative new semantics introduced by Hodges [3, 4]. This new semantics, called *team semantics*, defines the satisfaction relation with respect to *sets* of assignments (called *teams*), instead of single assignments as in the standard Tarskian semantics of first-order logic. Teams can be easily conceived as tables or data sets. The flexible and multidisciplinary interpretations of teams results in a rapid development of applications of the logics in recent years.

Formalization of Matthew effects. In this work, we describe three distinct types of Matthew effect, namely *direct Matthew effect*, *mediated Matthew effect* and *complete Matthew effect*, and we give formal definitions for them via independence logic. Consider the signature $\mathscr{L}$ that contains the equality symbol $=$, the constant symbols r for each real number $r \in \mathbb{R}$, the function symbols $+, -, \cdot, \div, (\cdot)^r$ for each $r \in \mathbb{R}$, relation symbols $\leq, \geq, <, >$ and other relevant non-logical symbols. We assume that the context of the Matthew effects in question is captured by a first-order $\mathscr{L}$-model M. The domain of an intended model M of a Matthew effect scenario consists of the set of all possible values of all data sets (e.g. real numbers, names of products, names of artists, etc.).

Given a data set and a system of equations that corresponds to a statistical analysis of the data set. We use $x, y, w, \dots$ to denote the variables in the data set, and we reserve the letter t for the time variable. We write $x_{(t)}$ for the value of the variable x at time t. For the formal definitions, following [7], we view the properties being defined as new atomic formulas and only give their corresponding team semantics.

- y is **(positively) dependent on** x, denoted $x \nearrow y\,|^{t}_{\vec{w}}$, if there exists an equation in the system such that for some threshold $\gamma \in \mathbb{R}$, $x_{(t-1)} \geq \gamma \Longrightarrow y_{(t)} = \beta_0 + \delta x_{(t-1)} + \beta_1 w_{1\,(t-1)} + \cdots + \beta_m w_{m\,(t-1)} + \epsilon$, where $\delta > 0$ is a constant, $w_1, \dots, w_m$ are dependent variables, $\beta_0, \beta_1, \dots, \beta_m$ are nonzero parameters and ϵ is an error term. In other words, if $x \nearrow y\,|^{t}_{\vec{w}}$, then, after x reaches the threshold γ, when all the other relevant variables $\vec{w}$ are held constant, we have $y_{(t)} - y_{(t-1)} = \delta \cdot (x_{(t-1)} - x_{(t-2)})$ for some $\delta > 0$. Formally, we introduce a new atomic formula $x \nearrow y \mid^t_{\beta,\gamma,\vec{w}}$, and define $x \nearrow y \mid^t_{\vec{w}} := \exists^1 \beta \exists^1 \gamma (\beta > 0 \wedge x \nearrow y \mid^t_{\beta,\gamma,\vec{w}})$.

- y is subject to a **(positive) direct Matthew effect**, denoted $\mathsf{ME}y \mid^t_{\vec{w}}$, if y is positively dependent on itself. Formally, we define $\mathsf{ME}y \mid^t_{\vec{w}} := \exists^1 \delta \exists^1 \gamma (\delta > 0 \wedge y \nearrow y \mid^t_{\delta,\gamma,\vec{w}})$.
- y is subject to a **(positive) x-mediated Matthew effect**, denoted $\mathsf{MME}y(x) \mid^t_{\vec{w}}$, if after some threshold γ, x is positively dependent on y, and y is positively dependent on x. Formally, define $\mathsf{MME}y(x) \mid^t_{\vec{w}} := \exists^1 \delta_1 \exists^1 \delta_2 \exists^1 \gamma_1 \exists^1 \gamma_2 (\delta_1 > 0 \wedge \delta_2 > 0 \wedge (y \nearrow x \mid^t_{\delta_1,\gamma_1,\vec{w}}) \wedge (x \nearrow y \mid^t_{\delta_2,\gamma_2,\vec{w}}))$.
- x and y are subjects to a **(positive) complete Matthew effect**, denoted $\mathsf{CME}(x,y) \mid^t_{\vec{w}}$, if y is subject to a positive x-mediated Matthew effect, y is subject to a positive direct Matthew effect, and x is subject to a positive direct Matthew effect. Formally, define $\mathsf{CME}(x,y) \mid^t_{\vec{w}} := \mathsf{MME}x(y) \mid^t_{\vec{w}} \wedge \mathsf{ME}x \mid^t_{\vec{w}} \wedge \mathsf{ME}y \mid^t_{\vec{w}}$.

Results. It is clear from its defining clause of team semantics that the auxiliary new atomic formula $x \nearrow y \mid^t_{\beta,\gamma,\vec{w}}$ we introduced is a Π_1 atom in the sense of [7], and therefore both definable and negatable in $\mathcal{I}$. Since first-order atomic formulas are negatable in $\mathcal{I}$ and the class of negatable formulas of $\mathcal{I}$ is closed under $\wedge$ and $\exists^1$ [7], we conclude that the formula $x \nearrow y \mid^t_{\vec{w}} = \exists^1 \delta \exists^1 \gamma (\delta > 0 \wedge x \nearrow y \mid^t_{\delta,\gamma,\vec{w}})$ is negatable and definable in $\mathcal{I}$. Similarly, the defining formulas $\mathsf{ME}y \mid^t_{\vec{w}}$, $\mathsf{MME}y(x) \mid^t_{\vec{w}}$ and $\mathsf{CME}(x,y) \mid^t_{\vec{w}}$ of the different types of Matthew effects are all definable and negatable in $\mathcal{I}$. This means that the completeness theorem of independence logic applies to the formulas defining different Matthew effects, and therefore many properties of Matthew effects can be derived formally in the system of [7]. Simple examples of such properties include: $\mathsf{MME}y(x) \mid^t_{\vec{w}} \vdash \mathsf{MME}x(y) \mid^t_{\vec{w}}$ (mediated Matthew effects are always *reciprocal* for the two variables involved) and $\mathsf{ME}y \mid^t_{\vec{w}} \vdash \mathsf{MME}y(y) \mid^t_{\vec{w}}$ (a direct Matthew effect is a special case of the mediated Matthew effect). More interesting properties will be explored in our future work.

Further research. Future research will be directed at formalizing the Matthew effect in a real-world context [1]. The authors analyze differentials in the recognition received by U.S. biomedical scientists who are awarded the prestigious Howard Hughes Medical Institute (HHMI) appointment, relative to scientists of comparable quality who are not awarded the HHMI affiliation. The empirical analysis reveals that HHMI-appointed scientists tend to earn greater recognition, especially if there is greater uncertainty about the quality of their output. This suggests important boundary conditions to the Matthew effect, which will be taken into account in our formal approach.

References

[1] P. Azoulay, T. Stuart, and Y. Wang. Matthew: Effect or fable? *Management Science*, 60(1):92–109, 2013.

[2] E. Grädel and J. Väänänen. Dependence and independence. *Studia Logica*, 101(2):399–410, 2013.

[3] W. Hodges. Compositional semantics for a language of imperfect information. *Logic Journal of the IGPL*, 5:539–563, 1997.

[4] W. Hodges. Some strange quantifiers. In *Structures in Logic and Computer Science: A Selection of Essays in Honor of A. Ehrenfeucht*, LNCS 1261, pp. 51–65. Springer, 1997.

[5] R. Merton. The Matthew effect in science. *Science*, 159(3810):56–63, 1968.

[6] J. Väänänen. *Dependence Logic: A New Approach to Independence Friendly Logic*. Cambridge University Press, 2007.

[7] F. Yang. Negation and partial axiomatizations of dependence and independence logic revisited. In *WoLLIC 2016*, LNCS 9803, pp. 410–431. Springer-Verlag.