Vistas from a drop of water

Johan van Benthem

"From a drop of water a logician could infer the possibility of an Atlantic or a Niagara without having seen or heard of one or the other." Arthur Conan Doyle.

1 What can we learn from a single fact?

Can logicians infer a Niagara from a single drop of water, as Conan Doyle said? We only observe singular facts. Can we draw any general conclusions from these? Logicians might give answers like this. We can never infer general laws from singular observations, though they can refute general laws, bringing down the mighty. Or to be more precise, the only general things we can infer from single facts are pompous reformulations, such as in the valid inference from Pa to $\forall x\ (x = a \rightarrow Px)$.

But this cannot be the whole story. The theme of inferring further new things from concrete instances arises in many places. Consider the 'problem of Locke-Berkeley' as discussed in Beth 1959: how can we infer general geometrical statements from the contemplation of a single triangle? Beth himself thought there was a confusion here that is easily dispelled by understanding how the rule of Universal Generalization functions in logical proof. But the debate about 'arbitrary objects', 'generic structures' and other ways in which observing typical objects can yield generality, both in daily life and in science, continues – and logic can even inform us about how much generality can be extracted from what (van Benthem 1981, Fine 1985).

In this piece, I will ignore the generic object perspective, interesting though it is, and focus on facts about ordinary situations and what these can tell us in general.

To start at base level, consider the simplest pattern of inference by *analogy*:

from Px, $x \sim y$ *to* Py.

This is perhaps the most frequent occurrence of reasoning in practice, with long historical credentials in many logical traditions. If P holds of object or situation x, and x stands in some suitable relationship to y, we conclude that P holds of y. But what is this relation $\sim$, and does standard logic endorse this sort of inference?

For a first illustration, let us simplify Conan Doyle's ambition to just the following. If x is a drop of water with property P, what objects y are sufficiently connected to x to allow for the conclusion Py? Various candidates come to mind. Topologically, drops of water are even isomorphic to the whole space they live in, so the topological structure we can see in the small holds in the large. [1] But of course, not everything goes: specific metric assertions about the size of the drop will not transfer.

From this brief and simplified formulation of the raindrop example, we can take two points. Transfer is possible when there is enough *similarity* between situations, where similarity can be defined in precise mathematical terms. But crucially also: what can be transferred depends on our *language* for describing properties. We will discuss this combined perspective in Section 2 below.

Our second example comes from ancient Indian logic (Bocheński 1961). Simplifying a bit, we are at the foot of a mountain, a situation from where we would like to draw inferences about what is the case on top of the mountain, a situation that is inaccessible to us. We see smoke, and we conclude that there is a fire on top of the mountain, using the connecting statement that "smoke means fire". Examples like this are quite persistent in the history of logic, witness this version from Moist logic in ancient China (Liu & Zhang 2007). You see an object in a dark room, but not its color. You see an object outside that is white. Someone tells you the two objects have the same color. You conclude that the object inside the room is white. [2]

Similar examples with transfer from actual situations to inaccessible situations occur right in modern times in situation theory (Barwise & Seligman 1995).

Examples like this do not turn on similarity between the two situations at issue. Consider again the above inference pattern

from Px, $x \sim y$ *to* Py.

[1] The same analogy inference even holds for a wide range of suitably stated material properties if one believes in the well-known principle of Homogeneity for the physical universe.

[2] The Chinese example was meant to highlight the three types of information coming together here: observation, communication, and inference – but what matters for us here is the transfer pattern.

This time the relation ~ is different. What matters to the information flow and its admissible inferences in the second type of scenarios is *correlation* of facts in situations, or in more general settings, between the behavior of the situations over time. This is a serious alternative perspective that we will discuss in Section 3 below.

But there is a further relevant distinction to be made. The preceding two perspectives on generality might be called static, focusing on *what* is the case. The situation we have access to satisfies some fact, there is a connection with another situation, the fact (or some suitably adapted variant of it) also holds there. This transfer is automatic. But there is also another, equally ubiquitous sense of learning from truth in single instances, where we have to do serious work, by reflecting on the *how*. We identify the *reason* for our saying that some fact holds in a particular situation, and then see, by reflecting on this reason, that it applies more generally.

For a concrete illustration, consider standard game-theoretic semantics for first-order logic (cf. the survey in van Benthem 2014). A first-order formula φ is true in model $\boldsymbol{M}$ with variable assignment s iff Verifier has a *winning strategy* against Falsifier in the evaluation game for φ in $(\boldsymbol{M}, s)$. These winning strategies (there can be more than one) are not just truth values: different winning strategies stand for different reasons why φ is true in $\boldsymbol{M}$ – and described in suitable generic terms, they can be played elsewhere, allowing us to see φ that holds in other models as well.

Now this brings us to a potential controversy. In the above terms, *model checking*, testing whether a given formula φ is true in a given finite model $\boldsymbol{M}$, starts looking like *proof*, showing that φ follows from premises Π identified in the process of model checking. But these things are very different: a statement $\Pi \models \varphi$ tells us something about *all* models of Π. And this difference also shows up in computational complexity. For first-order logic, model checking is decidable, but testing for validity is undecidable. So, our jump to generality seems to disregard a serious inevitable barrier in complexity: and as we all know, miracles do not happen in logic. [3]

[3] Model checking and testing for validity do coincide for suitably weak sublanguages of first-order logic admitting minimal models (Kolaitis & ten Cate 2015). Also, model checking, validity testing and model construction get entangled in practical tasks (Goranko 2018). We ignore this line here.

Even so, the interplay of semantic truth and proof is natural, and we will take it on board as a running theme in discussing transfer from the particular to the general.

It is easy to identify further ways in which singular facts can lead to general conclusions, especially if we also consider inductive and abductive reasoning. But we will not go there: our aim with this article is not to exhaust, but instruct and alert.

2 Similarity and preservation

Many relations between models support transfer of properties expressed in some matching language. A typical instance in logic textbooks is the

> *Isomorphism Lemma* If f is an isomorphism from model $\boldsymbol{M}$ to $\boldsymbol{N}$,
> then, for all first-order formulas φ, $\boldsymbol{M}, s \models \varphi$ iff $\boldsymbol{N}, f \bullet s \models \varphi$,
> where $f \bullet s$ is the assignment that sends variables x to $f(s(x))$.

In fact, all standard logical systems have this property, which is therefore one of the basic defining features of logical systems in Abstract Model Theory.

But isomorphic images of given models are not all that interesting, since they are really just presentations of the very same structure in different guise. Transfer gets more interesting when we weaken the language, and concomitantly, coarsen the structural relation between models. For instance, *modal formulas* are invariant for *bisimulation*, a much less demanding notion of structural similarity than isomorphism, which connects a given model $(\boldsymbol{M}, s)$ to many more variants. And there are many further similarity relations, each coming with their own special syntax for the invariant structural properties (cf. van Benthem 2018 for a survey). [4]

Summarizing all of this, picture a single model as a radio transmitter in a vast universe. Its truths transfer in circles around it, where larger circles represent weaker similarity relations, and what gets through gets ever less detailed, requiring ever more restricted, less expressive syntax. This broadcast metaphor supports generality from individual facts, of the first kind discussed in our introduction.

---

[4] We can even make the similarity dependent on the particular formula φ we started with, demanding only structural equivalence of models up to the quantifier depth of φ, by playing Ehrenfeucht-Fraïssé games up to a fixed finite length, or yet finer pebble versions thereof (van Benthem 2014).

But this perspective can still be generalized. So far, we considered *equivalences* of truth between models, but this is not needed at all. The standard *preservation theorems* of logical model theory tell us that special syntax supports unidirectional generality. For instance, a positive sentence true in a model ***M*** will still hold in all homomorphic images of ***M***, and many further such preservation results exist.

But going yet one step further, there is no need to just have the same formula transferred to other situations, witness the following generalized notion of consequence proposed in Barwise & van Benthem 1999. Let *R* be any binary relation between models (such as isomorphism, bisimulation, homomorphic image, submodel, ...).

Formula φ *entails* ψ *along R* if, whenever ***M*** *R* ***N***, and ***M*** $|= \varphi$, then ***N*** $|= \psi$.

One motivation for this goes back to our earlier examples: we observe φ in our actual situation ***M*** to learn that ψ in some inaccessible (but still related) situation ***N***. Standard logical consequence is then the special, somewhat timid, case of inference where *R* is the identity relation. Entailment along relations has interesting properties, such as supporting new types of interpolation theorem for logical systems, but these need not detain us here. Our main point is that generality from similarity and transfer is entirely feasible, when we keep a clear view of the balance involved between similarity relations and the syntax of the observed individual facts.

The above discussion of the broadcast setting naturally shifted to consequence relations, and thus, it also suggests a proof-theoretic perspective. By the completeness theorems for many logics, there is a proof system for transfer consequences if the model relation can be defined in the language of the logic. Whether this is possible depends. Definability holds for transfer under isomorphism and homomorphism in first-order logic, but the crucial modal invariance of bisimulation is not definable inside the modal language (van Benthem, ten Cate & Väänänen 2009). [5]

A particular case of definable cross-model relations arises when we think of models as situations that can be changed, for instance, by update with new information or

[5] Axiomatizing complete meta-model-theories in this style may have its surprises. E.g., the modal logic of bisimulation relating modal models may be undecidable: the back-and-forth property of bisimulation defines a grid whose complete modal logic is known to be of high complexity.

other actions. In that case, the discussion of learning from single situations acquires a new flavor, entangled with proof, that we will discuss in Section 4.

3 Correlation

Next, let us return to the Smoke and Fire example of our Introduction. Here some different logical structures emerge. In the simplest case, we have two situations s_1 and s_2, one close to us, one far away: there is a fact p about s_1 that we can observe, and since s_1 and s_2 are correlated, this tells us that some other fact q holds in s_1.

To model this, we need to think about the logic of *correlations* (cf. van Benthem & Martinez 2008, a wide-ranging discussion of the variety of notions of information found in contemporary logic). We have a possibly large set of situations or locations that can have various properties, and there may be constraints on the occurrence of these: in the simplest case, as equivalences $p{:}\ s_1 \leftrightarrow q{:}\ s_2$. In this setting, transfer is not by similarity, but by constraints or correlations among situations.

Correlations can have many sources: from ontological (say, through laws of nature) to conventional (say, notes played by instruments in a performing a piece of music). These sources are not themselves logical, but there are interesting logical issues in understanding correlation. The above equivalence $p{:}\ s_1 \leftrightarrow q{:}\ s_2$ may seem just brute force stipulation, but more can be said when we enrich the setting a bit.

Consider a system of many situations evolving through time, where each situation can have or lose properties *p, q,* ... This results in behaviors, histories whose stages are truth-value assignments to *p, q, ...* at each situation. [6] Correlation now means that not all histories are possible: there are *gaps*. These gaps encode important information, namely that there can be various dependencies. Say, if we fix the value of p among the admissible histories, we automatically also fix the value of q – or, if we try to change the value of p at one location inside an admissible histories, we find that that of q at another location must change as well. This is just one case, many further natural forms of dependence can occur in a given set of behaviors.

In a stark mathematical model, we can just think of the preceding behaviors as assignments of values to variables, and what we end up with are models for first-

[6] Significantly, this is again a move to dynamics and time, which will occur again in Section 4.

order logic that do not have the full space of all functions from variables to objects as available assignments, but only a subset. These models validate a decidable version of first-order logic, the basic logic that remains after the standard Tarskian assumption of total independence of variables has been lifted (van Benthem 1996). [7] Extensions of the standard first-order language with explicit information about dependencies still validate a decidable base logic (Baltag & van Benthem 2018).

Of course, it will not always be the case that knowing the value of one variable tells us the exact value of another. There can be weaker dependencies where restrictions on values of *x* merely induce restrictions on values of *y*. But these, too, may convey general information out of particular observations in our sense.

Logical languages and systems accessing this correlation structure occur in a large variety. These range from decidable first-order modal-style logics for dependence and independence (Andréka, van Benthem & Németi 1998, Wang 2016, Baltag & van Benthem 2018) to second-order dependence logics in the style of Väänänen 2007. In whatever format, such logics can be seen as vehicles for explaining how general information can come from observations of a single situation or variable.

In this setting, generality from facts about single situations does not arise from similarity, and it does not broadcast through the whole meta-universe of all models. The generality rather arises from constraints encoded *inside the current model* of our system of situations, our 'distributed system' in the sense of Barwise & Seligman 1995, and its universal spread from one situation to others is confined to the situations (or more abstractly, variables) represented inside that model.

Instead of our earlier broadcast metaphor for similarity, one can think here of a system of linkages through information channels – or more irrelevantly, of an informational puppet theatre with rods and strings between puppets and players.

This style of viewing things through linkages has great power, it applies very widely, and its potential for systematic theorizing has not been exhausted by far.

[7] The lower complexity is no accident. Andréka, van Benthem, Bezhanishvili & Németi 2014 discuss general assignment models as Henkin models, and clarify connections with algebraic semantics.

At this point, one might ask whether the similarity view and the correlation view are related: could not one subsume the other? In special cases, one can, but I do not see much advantage in undermining an illuminating conceptual distinction. [8]

4 Interfacing models and proofs: dynamic-epistemic update

In this section, we show the *how* perspective of our introduction in action, by discussing a concrete form of information flow where single models meet with proof.

A particular case of using single models can be found in the dynamic-epistemic logic of information-driven agency (van Benthem 2011). Consider the by now stan-dard example of the Three Cards. Cards *red, white, blue* are dealt to players: *1, 2, 3,* one for each. Initially, each player sees his own card only. The real distribution over *1, 2, 3* is *<red, white, blue>*. Now player *2* asks player *1*: "Do you have the blue card?" Next, *1* answers truthfully "No". Who knows what then?

Here is the effect in words:

> Assuming the question is sincere, *2* indicates that she does not know the answer, and so she cannot have the blue card. This tells *1* at once what the deal was. But *3* does not learn, since he already knew that *2* does not have blue. When *1* says she does not have blue, this now tells *2* the deal. *3* still does not know even then. But since 3 can go through the above reasoning, he knows that the others know.

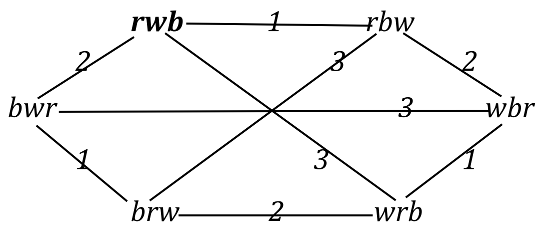
Now let us turn to models, the vehicle for our discussion so far. It is standard to picture the information flow in this scenario by means of updates in a diagram, making these considerations geometrically transparent. [9]

The initial situation can be represented as the following epistemic model, with the actual deal of the cards marked in bold-face as ***rwb***, and the indexed lines indicating situations that the players cannot distinguish visually:

---

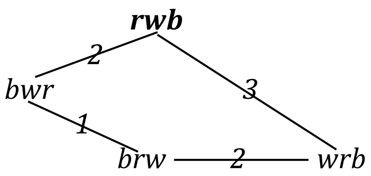
[8] Perhaps a more fruitful approach to analyzing transfer is trying to merge the two perspectives. If one were to include universes of models as in Section 2, correlation would again have to arise from gaps. We might drop some models from the relevant family of models around our model of origin, constraining the domain and range of familiar operations on, or relations between models. The effect of this move, reminiscent of 'protocol models' for dynamic-epistemic logic (van Benthem, Gerbrandy, Hoshi & Pacuit 2009), on exploring meta-model-theory remains to be explored.

[9] Technically, these diagrams are models for epistemic logic, but details need not concern us here.

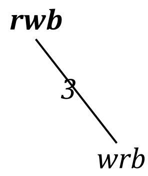


Here are the effects of the successive updates:

After *2*'s question:



After *1*'s answer:



We see in the final diagram that players *1, 2* know the initial deal ***rwb***, as they have no uncertainty lines left. But *3* still does not know, given her remaining line, but she does know that *1, 2* know – and in fact, the latter fact is common knowledge. Here updates with new information work as follows. Publicly learning that a statement φ is the case restricts the current model $(\boldsymbol{M}, s)$ (with actual world s) to the definable submodel $(\boldsymbol{M}/\varphi, s)$ whose worlds are only those worlds in $\boldsymbol{M}$ that satisfy φ.

Now here is an issue. The graphical representation just given, though concrete and in line with diagrams that people draw naturally, is just about one single situation. But one could read the Three Cards scenario as being more general. It stipulates just some facts about the start of a game, and makes an assertion about what happens after two specific rounds of update. This should apply to any initial model satisfying the facts, so we are really aiming for a universal statement, and a proof for it.

Here is some syntax to make this precise. The initial conditions of the scenario are defined by a formula φ expressing the *precondition*, the final effects are defined by a *postcondition* ψ. Updates are triggered by events $!\alpha$, where α is the true information conveyed, of which there may be one or more. Let $\boldsymbol{\alpha}$ stand for the whole sequence. The above suggests that we want to pass on to valid general insights of the form

$$\varphi \to [\boldsymbol{\alpha}]\psi,$$

where $[\boldsymbol{\alpha}]$ is a dynamic modality stating what is the case after $\boldsymbol{\alpha}$ has been executed.

Now, the single model we have considered does not support these per se, but it does contain clues for a more general analysis. Analyzing the update operation $!\alpha$ used above, we see that it validates a number of 'recursion laws' that describe effects of an update in terms of what was true one step before. The most striking of these is the recursion law for knowledge of agents after update, which reads as follows:

$$[!\alpha]K\varphi \leftrightarrow (\alpha \rightarrow K(\alpha \rightarrow [!\alpha]\varphi)).$$

For the complete logic of information update, we refer to the cited literature.

The key point for us is just this. Using these principles backwards, given a post-condition ψ and update sequence $\boldsymbol{\alpha}$, we can compute in a stepwise manner just what is needed in the initial situation for the update sequence to have the intended effect. [10]

Thus, analyzing our example reveals a general pattern. Any model satisfying a pre-condition that implies the condition just computed will have the required effects. [11]

This case study of information update shows how analysis of single models and their definable updates can go hand in hand with proof analysis, where analyzing well-chosen single models can suggest useful proof-theoretic principles which can then be used to extract more generality from the initial scenario.

Of course, as stated in Section 1, this does involve a shift from the broadcast scenario of Section 2, where transfer comes for free along similarity relations. One has to do real work to uncover the proof-theoretic recursion laws that induce generality.

Finally, in practice, entanglement of models and proofs is widespread. Solving real-life practical problems often involves an interplay of model-checking over diagrams,

---

[10] Technically, this procedure computes 'weakest preconditions' for the intended effect. One might also think that the Three Cards asks for a 'strongest postcondition' given just the precondition φ and the update sequence $\boldsymbol{\alpha}$, but computing this is a much more delicate matter (van Benthem 2011).

[11] Sometimes, single models are all that is needed. The preconditions in the original scenario may be so strong that they admit of only one model up to bisimulation. Arguably, this is true in the Three Card scenario and other famous update puzzles such as Muddy Children. Then updates on this single model capture the scenario as stated, up to some inevitable bisimulation variants. But the pre-condition-postcondition analysis still has its uses, as it will apply to changes in the original scenario.

when useful and fast, with symbolic proof steps that agents already know to bypass more tedious episodes of model-checking (van Benthem 1996). [12]

5 Conclusion

We have discussed several roads to generality from one single situation, all of them with a solid basis in logic. There is similarity and laws of transfer, there is correlation and dependence logics, and orthogonally to this, there is the interplay between seeing individual facts and reflecting on the proofs that led to their recognition. And these roads co-exist, there is absolutely no need to prefer one over the other.

Here is one further thought on how to extract generality. Either we already *have* it, as we saw with similarity and correlation, or we can *achieve* it by reflecting on specific truths and moving to suitable abstractions, perhaps resulting in proof systems that capture part of our handling the models. The latter procedure is not deterministic, and to some extent, it may be an art – but it is an important creative art. [13]

We end with one more thought on this process of abstraction from single cases. The vehicle for our discussion was logic and its formalisms. However, when describing what is general about concrete situations, logical syntax sometimes seems too rigid and particular, and *natural language* has the edge. For instance, in agent scenarios like the Three Cards puzzle, standard logical formulas force us to be specific about small details, whereas natural language has wonderfully concise indexical expressions such as "each agent knows his own card, but not those of the others". It is easy to find further examples of this virtue. Even mathematicians use natural language when describing high-level features of models or proofs in an illuminating manner. Explaining the fascinating interplay of natural and formal languages in achieving generality is not my task here, [14] but it deserves at least this much mention.

---

[12] Even more subtle mixtures of proof and model checking occur in efficient symbolic model-checking techniques for dynamic-epistemic logic, cf. van Benthem, van Eijck, Gattinger & Su 2018.

[13] It may seems as if this article suggested that, to achieve abstraction and generality, we must use logical syntax. But as explained in the influential manifesto Halpern & Vardi 1991 that advertized the virtues of model checking over proof in concrete computational applications, the art is rather what to leave in the meta-language of the models and what to highlight in logical syntax.

[14] As a slightly tongue-in-cheek example, consider the phenomenon of *ambiguity*, usually considered a disadvantage, but in fact a stimulus for creativity. Take the very metaphor of the "drop

Thus, our discussion of moving from the small to the large comes to an open end.

6 References

H. Andréka, J. van Benthem & I. Németi, 1998, 'Modal Logics and Bounded Fragments of Predicate Logic', *Journal of Philosophical Logic* 27:3, 217–274.

H. Andréka, J. van Benthem, N. Bezhanishvili & I. Németi, 2014, 'Changing a Semantics: Opportunism or Courage?', in M. Manzano, I. Sain and E. Alonso, eds., *The Life and Work of Leon Henkin*, Birkhaüser Verlag, 307–337.

A. Baltag & J. van Benthem, 2018, 'A Decidable Logic of Functional Dependency', working paper, ILLC, University of Amsterdam.

J. Barwise & J. van Benthem, 1999, 'Interpolation, Preservation, and Pebble Games', *Journal of Symbolic Logic* 64:2, 881–903.

J. Barwise & J. Seligman, 1995, *Information Flow*, Cambridge University Press, Cambridge UK.

J. van Benthem, 1981, 'Historische Vergissingen? Kanttekeningen bij de Fregeaanse Revolutie in de Logica', *Kennis en Methode* 5:2, 94–116.

J. van Benthem, 1996, *Exploring Logical Dynamics*, CSLI Publications, Stanford.

J. van Benthem, 2011, *Logical Dynamics of Information and Interaction*, Cambridge University Press, Cambridge UK.

J. van Benthem, 2014, *Logic in Games*, The MIT Press, Cambridge MA.

J. van Benthem, 2018, 'Semantic Perspectives in Logic', to appear in G. Sagi & J. Woods, eds., *The Semantic Conception of Logic*, Cambridge University Press, Cambridge UK.

J. van Benthem, B. ten Cate & J. Väänänen, 2009, 'Lindström Theorems for Fragments of First–Order Logic', *Logical Methods in Computer Science* 5:3, 1–27.

J. van Benthem, J. van Eijck, M. Gattinger & K. Su, 2018, 'Symbolic Model-Checking for Dynamic–Epistemic Logic', *Journal of Logic and Computation*, *28*:2, 367-402.

of water" which opened this article. This expression has a slight ambiguity, since 'drop of water' might also suggests a waterfall, as in dropping some object. But then we are much closer to the Niagara, and in fact also to the dynamic perspectives discussed later on in this article.

J. van Benthem, J. Gerbrandy, T. Hoshi & E. Pacuit, 2009, 'Merging Frameworks for Interaction', *Journal of Philosophical Logic* 38:5, 491–526.

J. van Benthem & M-C Martinez, 2008, 'The Logical Stories of Information', In P. Adriaans & J. van Benthem, eds., *Handbook of the Philosophy of Information*, Elsevier Amsterdam, 217–280.

E.W. Beth, 1959, *The Foundations of Mathematics*, North-Holland, Amsterdam.

I. M. Bocheński, 1961, *A History of Formal Logic*, Notre Dame Press.

B. ten Cate & Ph. Kolaitis, 2015, 'Schema Mappings. A Case of Logical Dynamics in Database Theory', in A. Baltag & S. Smets, eds., *Johan van Benthem on Logic and Information Dynamics*, Outstanding Logicians Series, vol., Springer, Dordrecht, 67–100.

K. Fine, 1985, *Reasoning with Arbitrary Objects*, Blackwell, London.

V. Goranko, 2018, 'Model Checking and Model Synthesis from Partial Models', working paper, Philosophical Institute, Stockholm University.

J. Halpern & M. Vardi, 1991, 'Model Checking versus Theorem Proving: A Manifesto', in *Artificial Intelligence and Mathematical theory of Computation*, Academic Press Professional, San Diego, 151–176.

J. Väänänen, 2007, *Dependence Logic*, Cambridge University Press, Cambridge UK.

Y. Wang, 2016, 'Beyond Knowing That: A New Generation of Epistemic Logics', Department of Philosophy, Peking University.

J. Zhang & F. Liu, 2007, 'Some Thoughts on Mohist Logic', in J. van Benthem, S. Ju & F. Veltman, eds., *A Meeting of the Minds: Proceedings First LORI Workshop on Logic, Rationality and Interaction*, College Publications, London, 85–102.