The Stories of Logic and Information DRAFT

Johan van Benthem & Maricarmen Martinez, March 2007

Commentator: John Perry

Introduction Information is a notion with wide intuitive appeal, and many formal paradigms claim part of it, from Shannon information theory to Kolmogorov complexity. Information is also a widely used term in logic, but a similar diversity repeats itself: there are several competing logical accounts of this notion, ranging from semantic to syntactic. In this chapter, we will discuss three major logical accounts of information.

The first is associated with sets of possible worlds (taken in a relaxed methodological sense) in epistemic logic, and we call it *information as range*. This setting brings its own agenda of what information is primarily good for and the further notions it is associated with. These include epistemic attitudes of agents toward brute information, such as knowledge or belief. Also, from the start, epistemic logic focuses on different information for different agents about the facts, and about each other, and what happens when these varieties meet and merge. The latter theme leads to another key issue for information in general, and inside logic: information is always to be understood in connection with some dynamic process using and transforming it. We will look at *dynamic logics of information transfer* for this purpose, which form a natural calculus of changes in information ranges triggered by events of observation or communication.

The second major logical strand in the notion of information high-lights connections between the components of a distributed system whose parts show dependencies. This notion of *information as correlation* has been developed in situation theory, starting from a naturalist theory of meaning for agents living in information-rich physical environments. The correlation paradigm brings with it a further agenda of the structure of situations –well-founded or 'circular'–, ways of classifying them, constraints on their joint behaviour, and channels which allow for information flow. Information as range and information as correlation are compatible semantic notions, and we show how they can be merged, in the spirit of modern logics of dependence and interaction.

Next, we move to a third major logical sense of information, oriented toward syntax, symbol manipulation, and computation. Thinking of information as encoded in sentences at some abstraction level, we then come to the idea of *information as code*. In this more concrete combinatorial setting, the major dynamic processing paradigm is 'inference' in some suitably abstract sense, and the relevant logical sub-discipline is no longer model theory, but *proof theory* and related modern theories of abstract computation.

Semantic and syntactic perspectives have always co-existed in logic, and their interplay is at the heart of the celebrated *completeness theorems*. We will show how this harmony also links our different logical views of information. Finally, we touch upon some other topics, such as the increasing role of non-linguistic visual information carriers in logic, as well as the border-line with quantitative approaches such as probability theory.

After this top-level sketch of what we are after, let's plunge into concrete logical matters.

1 Information in Logic

Information is a notion with a somewhat precarious status in logic. One often uses it colloquially to explain to beginning students what logic is all about, and a favourite metaphor is then that deduction is useful for 'extracting information' from the data at our disposal. Say, you know that $A \vee B$ and you learn that $\neg A$. Then logic tells you that you now know B, since the following basic inference schema is valid:

$$A \vee B, \neg A / \neg B.$$

By the way, this is also the schema behind *Sudoku's* and other logic puzzles sweeping the planet – so the reality of these phenomena of learning through logic is well-attested.

But what one metaphor gives, another one takes. Further on in the course, we also cheerfully tell students that the hall-mark of logical validity is its 'analyticity'. And that says that the conclusions 'do not add information' to what was already there in the premises. Indeed, there are many different views of information in logic, all of them implicit. The field has official definitions for its central concepts of proof, computation, truth, or definability, but not of information! Indeed, many logicians will feel that this is significant. We do not need this notion in the mechanics or even the foundations of the formal theory – or, as Laplace said to Napoleon, who inquired into the absence of God in his *Mécanique Céleste*: "Sire, je n'avais pas besoin de cette hypothèse".

Information, proof, and computation Still, more can be said. There are several areas in modern logic where notions of information emerge naturally. Indeed, logic can be seen as an Information Theory with impact across many disciplines. So far, we have already noticed the connection between information and inference, i.e., the *proof-theoretic stance toward logical validity*. Information states are then stages in a deduction, and informational moves are proof steps, or more general computation steps. For a concrete illustration, think of successive stages in the solution of a 3x3 Sudoku:

1	•	•
•	•	*2*
•	•	•

,

1	•	*3*
•	•	*2*
•	•	•

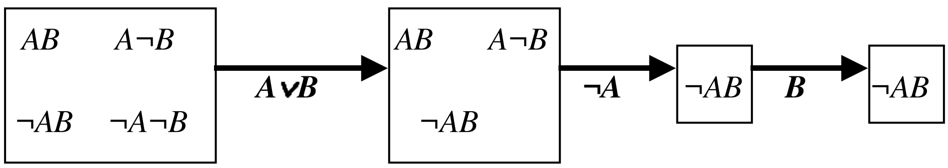
,

1	*2*	*3*
•	•	*2*
•	•	•

...

Each successive diagram displays a bit more information about the eventual solution. Thus, on this view, information is brought to light in logical proof theory or the theory of computation, and inferences are actions transforming the current state.

Information range and update To find further logical notions of information, however, consider just the simplest story behind the above inference scenario. Premises do not drop out of the sky! You are in a café with a friend, who has ordered Applejack, while you took a Brandy. A new waiter comes back from the kitchen with two glasses. What we see around us every day is that the new waiter asks who has the Applejack (say), puts it down, and then puts the other glass without asking. This scenario shows two informational processes intertwined. The final inference by the waiter produces information in the earlier sense, but so did your answer to his question, which was crucial to the whole information flow, be it in another sense. Your response did not involve deductive steps on your or his part, but rather changes in information *through observation*, giving us new facts. For instance, before learning the premises of the earlier inference, your information state contains 4 options. Receiving the first premise $A \vee B$ cuts this down to 3, and receiving the second premise $\neg A$ cuts down to just one option: $\neg A$ & B is the actual situation. The rest is indeed deductive inference, since further update, say with the remaining fact B, would not change the final state any more:



Again there is a process of information flow, now through update. This time, the process is related to the *semantic view of validity*, which says that the conclusion must be true in all models for the premises. We do not know what the real situation is, but we have some information about it, encoded in a current *range of models*. And the relevant informational actions this time are updates: each subsequent assertion that we learn decreases that range, to a first approximation. This notion of information is found in epistemic logic and related semantic paradigms which we will discuss below.

Information and interaction As we said, the new information feeding deduction may be seen as coming in through observation in some general sense. But questions and answers also highlight a third paradigm for logical validity, which involves many actors, dialogue, and argumentation. This is the world of Plato's *Dialogues* and public debate in the Greek polis, rather than the abstract proof structure of Euclid's *Elements*. On this *pragmatic view of validity* we are after strategies which agents can use for winning debating games, or achieving other crucial tasks involving information and persuasion. This brings us to information as what flows in communication, and more generally, as the lubricant of interaction. We refer to [Chapters Devlin, Walliser] for more on this take, which fits well with current views of logic as a theory of intelligent interaction, with game theory as a natural ally. As with the proof-theoretic and semantic views, we find information arising in a process, not as an absolute commodity by itself.

Information and Correlation But there are still further perspectives in our logical panorama! The research program which put information on the map perhaps most forcefully as a major theme for logical study per se was *Situation Theory*. Here, flow of information is not seen primarily as driven by events of observation or communication between agents. It rather rests on persistent 'constraints' in the world, which correlate the behaviour of spatio-temporally different situations. Thus constraints provide *information channels* between situations, and we can avail ourselves of these to draw conclusions about what we may not be able to observe directly.

Again, this may be tied in with notions of inference – once we are willing to step away from the usual varieties. E.g., one recurrent 'syllogism' in the ancient Indian logic tradition runs roughly as follows [Staal]. I am standing at the foot of the mountain, and cannot inspect directly what is going on there. But I can make observations in my current situation. Then, an inference might work as follows:

> "I see smoke right here. Seeing smoke here indicates fire on the mountain.
> So, there is a fire on the mountain top."

This version is simplified: Indian syllogisms had other interesting non-Aristotelean features. Even so, compare this with the usual notion of inference, which is about one fixed situation. The main new idea will be clear: logical inference can also cross between situations. Given information channels between situations, observations about one may give reliable information concerning another. Indeed, the Indian example is almost the running example in Barwise & Seligman 1995 on seeing a flash-light on the mountain. Other *très* Indian examples include observing a coiled object in a dark room, and using logic, rather than touch, to find out if it is a piece of rope or a cobra.

The goal of this chapter One purpose of this chapter is bringing out information as a theme running all through logic, even when it is usually left implicit. The resulting take on logic today and some of its major features is shown in the following sections. This is mainly a matter of presentation and re-telling existing stories. But there are genuine conceptual problems to be addressed as well. The many notions of information in logic pose a problem because they are so different, and yet they each make some valid point.

The second purpose of this chapter is to provide a more unified perspective, pointing out some new connections, and raising some further questions. This sort of question has not been addressed in the philosophy of logic, as far as we know, where the agenda still seems largely concerned with somewhat fossilized questions from the past.

To see why there is a potential for integration, one can note at once that the different logical information theories have clear similarities. In particular, all seem to involve an interplay between what may be called *statics* and *dynamics*. There are background structures representing the information, but these only make sense as vehicles for various processes of information flow. Yet epistemic logic, situation theory, and proof theory all approach information dynamics from different stances. We think this diversity may be a blessing rather than a curse. We will elaborate on the stances, and discuss interactions – such as that between communication and correlation, or between observation and deduction. But we do not (yet) have one unified notion of information coming out of all this, making logic show the same diversity one finds with the notion of information in general. Indeed, we will also high-light some connections to adjoining fields.

Finally, similar issues also occur in other chapters of this Handbook. Just compare [Baltag et al.] on epistemic dynamics, [Rott] on belief revision, [Kamp & Stokhof] on linguistic communication, and [Abramsky] on the information flow in computation.

2 Information as range: state spaces, epistemic, and doxastic logic

In this section, we develop the notion of information as range. In our view, current epistemic logic (broadly understood) is the 'information theory' that goes with this. We show how, just as in other formal theories, this brings along an agenda of further themes showing how this information functions, how it can be computed with, and what basic methodological issues it gives rise to. The headings in our text identify what these are.

2.1 Information, sentences, and state spaces

Successive assertions inform us about a situation which the current discourse is about. There are two logic-based ways of thinking about growth of information in this setting. One is in terms of *syntax*. Assembling assertions over time creates an ever-growing 'book' of sentences, and the information is the set of all these – perhaps also including their inferential connections, and maybe even rankings as to relevance or plausibility. This syntactic view of the accumulated information at our disposal may be the most attractive from a computational point of view. But it is not the main focus of this chapter, except for Section 6 below which makes syntax and proof the locus of information.

Instead, here is the natural *semantic* picture of information growth. On this inverse view, extra information does not add things: it rather shrinks something, viz. the current range of options for what the real situation might be. This is what we saw with the earlier update pictures for the inference $A \vee B$, $\neg A$ / $\neg B$, where the initial state of ignorance had 4 possible options, of which 3 remained after the input $A \vee B$, and only 1 after the further input of the premise $\neg A$. The inverse relationship can be shown as follows, with T for sets of formulas, and $MOD(T)$ for the class of models making all of T true:

$$T \subseteq T' \text{ iff } MOD(T) \supseteq MOD(T')$$

Using sets of models as information ranges is close to the 'state spaces' which [Carnap] used for describing the information associated with an assertion.

Of course, sets of models are rather rough counterparts to sets of sentences, since $MOD(T) = MOD(T')$ for logically equivalent sets of assertions T, T', even when these are vastly different syntactically. To most logicians, this is a virtue, as they find 'details of syntax' irrelevant to content (as long as they are, one hopes, not reading love letters). Nevertheless, one can also apply finer sieves, and Carnap himself also used syntactic 'state descriptions' for models as the basis for computations in his inductive logic. Indeed, much of the discussion of 'propositions' and 'meanings' in the philosophical literature [Lewis, chapter Kamp & Stokhof] might be seen as the search for a level of information in between mere sets of models and every last detail of syntax.

In Sections 2, 3, 4, 5 we continue with the semantic view of information as range.

2.2 Knowledge and epistemic logic

The best-known paradigm incorporating information as semantic range is *epistemic logic* as proposed by Hintikka 1962, and developed by many authors since, across different disciplines such as philosophy, computer science, and economics. Its main ideas are easy to describe: the models describe information ranges for agents, while the matching language describes a notion of knowledge that might be paraphrased as "to the best of the agent's information". Here is how this works in more formal detail.

Logical basics The syntax has proposition letters $p, q, \ldots$, Boolean connectives $\neg$, $\vee$, and modal operators $K_i\phi$. The latter express that agent i *knows that* ϕ, while the dual $\langle i \rangle\phi = \neg K_i \neg\phi$ says that i *considers* ϕ *possible*. The following semantics provides a precise underpinning for this intuitive reading. Models $\boldsymbol{M}$ for the language are triples

$$(W, \{\sim_i \mid i \in G\}, V),$$

where W is a set of worlds, the $\sim_i$ are binary accessibility relations between worlds, and V is a propositional valuation. The worlds ('states', 'situations', ...) in the set W represent the options for how the actual situation might be, while the relations $\sim_i$ encode the uncertainty, or alternatively, the current information of the agents:

$x \sim_i y$ says that, at world x, i considers y an option for being the actual world.

These accessibility relations may be different for different agents, who need not all have the same information. One often takes the $\sim_i$ to be equivalence relations, but this is not crucial to epistemic logic in general (indeed, it validates some much-discussed features of agents like 'positive' and 'negative introspection' concerning one's own knowledge). Now the semantic truth condition for the knowledge operator makes the most evident stipulation in this setting, using a universal quantifier over the current information:

Agents know what is true throughout their current range of uncertainty:
$\boldsymbol{M}, s \models K_i\phi$ iff for all t with $s \sim_i t$: $\boldsymbol{M}, t \models \phi$

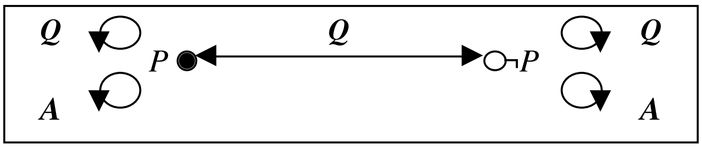
The dual $<i>\phi$ is then the existential quantifier 'in some currently accessible world'.

We follow the established 'knowledge' terminology for the operator $K_i\phi$ of epistemic logic in much of what follows. Even so, the more neutral and less ambitious term

'to the best of i's information' for $K_i\phi$

would better state what the above universal quantification over the current information range really achieves, and how epistemic logic serves as an information theory. We will briefly address this point again when discussing connections with formal epistemology.

Factual and higher-order information Simple as it looks, the epistemic language can formulate non-trivial scenarios. Consider a model for two agents ***Q***, ***A*** ('questioner' and 'answerer', as we shall see in a moment) with one world where the atomic fact P holds, and another where it fails. We assume that the real world (there always *is* one!) is the one indicated by the black dot, though this is an outsider's annotation, rather than the agents' own information. Labeled lines linking worlds indicate uncertainties. In particular, ***Q*** does not know which world is the actual one, while ***A*** is better informed: if the actual world has P then she knows that is the case, and if it does not, then she know that, too:



M

This diagram also encodes 'higher' information that agents have about their own and each others' information. For instance, in a sense to be defined below, ***Q*** knows that she does not know if P while ***A*** does. Hence, it would be a good idea for ***Q*** to *ask a question* to ***A***, and find out. But before getting to that, let us take stock of the model ***M***, representing the current state of the informational process. In formulas of the epistemic language, here are a few things which are true in the world to the left:

> *P*, K_AP, $\neg K_QP$, $K_Q\neg K_QP$, $K_Q(\neg K_QP \wedge \neg K_Q\neg P)$ (***Q*** knows that she does not know that *P*), $K_Q(K_AP \vee K_A\neg P)$ (***Q*** knows that ***A*** knows whether *P*)

Thus, this simple logical language for information can express complicated multi-agent patterns. In particular, *iterations* of knowledge about oneself and others reveal something essential about the notion of information as used by humans:

> *'Higher-order information' about the (lack of) information of ourselves and others seems just as important as ground-level factual information!*

Communication and information flow As in other information theories, such as Shannon's account of channel transmission, information really comes into its own only in a dynamic setting of *communication* and *information flow*. As a simplest case of this phenomenon in epistemic logic, consider the following conversational scenario:

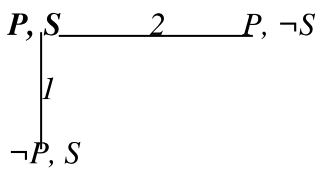
(a) ***Q*** asks ***A*** the question *"P?"*,
(b) ***A*** gives the true answer "Yes".

Then the situation depicted in the model ***M*** changes, since information will start to flow. If this episode is simple cooperative Gricean communication, the question (a) itself conveys the information that ***Q*** does not know the answer, but also, that she thinks ***A*** might know. In general, this can be highly informative and useful to the answerer. Next, the answering event (b) conveys that ***A*** knows that *P*, and its public announcement in the group *{**Q**, **A**}* makes sure that ***Q*** now also knows that *P*, that both agents know this about each other, and so on to very depth of iteration. In terms of Lewis 1969, following the episode, the agents have achieved *common knowledge* of *P*.

Epistemic logic describes many properties at the stages of this question–answer process: before, during, and after. In particular, it also has notions of *group knowledge*, doing justice to the fact that, when individual agents meet, one can and must also talk about knowledge possessed by groups of them. For instance, $C_G\phi$ ('ϕ is common knowledge in group *G*') is read in the above information models as follows:

> ***M***, $s \models C_G\phi$ iff for all worlds *t* reachable from *s* by some finite sequence of $\sim_i$ steps ($i \in G$): ***M***, $t \models \phi$.

Remark Incidentally, to show that knowledge about others' ignorance can be helpful in communication, consider the following model, with the actual world in bold-face:



Neither agent *1* nor *2* knows the real situation here. But if *2* were to say that she does not know, this is informative to *1*, since it rules out the bottom-most world. Hence, *1* would know the real situation, and could inform *2*. Alternatively, if both were to state *simultaneously* that they don't know, the actual situation would be reached at once. Thus again, contrary to received opinions in some circles, the simple semantics of epistemic logic can model, and indeed classify, quite interesting informational settings.

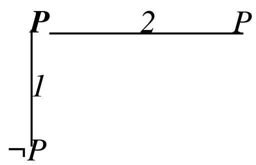
There is more to be said in logic about the information flow in these examples. We will look at the precise dynamics in Section 3 below. For the moment, we just note that formal definitions alone do not really show the power of a framework. Only a matching *art of modeling* will tell us about the real merits of a formal information theory. And indeed, in the hands of skilled practitioners, epistemic logic is an information theory which handles much more complex scenarios than the baby examples we give here.

Logic as information theory Viewing epistemic logic as an information theory on a par with others raises some immediate questions. What does the usual logical agenda mean in this light? Does it provide a calculus for information comparable to that of Shannon-style or Kolmogorov-style information theory? Most standard results in epistemic logic take one of the following forms. One looks at (a) expressive power and *definability*, (b) *axiomatic completeness* for proof systems, or (c) computational *complexity* of tasks like model checking or proof search. We will not survey all these issues in detail here, but just discuss some re-interpretations from an informational stance.

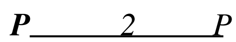
Digression: same information and invariance One crucial issue in logical languages is the harmony between the expressive power and matching *semantic invariances* between models. In particular, the latter style of thinking raises the conceptual issue of

When are two given information models the same?

We cannot just take things at face value, and assume that geometrically different diagrams automatically model different information. Consider the following variant $\boldsymbol{M_1}$ of the epistemic model in the above scenario, with the actual world again in bold-face:



Let 2 announce the true fact that she does not know that $\neg P$. The updated model is $\boldsymbol{M_2}$:



But clearly, $\boldsymbol{M_2}$ is essentially *the same* as the following one-world model $\boldsymbol{M_3}$:

P

as can be seen by identifying both worlds in $\boldsymbol{M_3}$ with the single one in $\boldsymbol{M_3}$. Indeed, understanding what makes different diagrams carry 'the same' information is a good test of our understanding of the notion of information itself. The answer given to this by epistemic logic is clear. Two information models *(**M**, s)* and *(**N**, t)* satisfy the same epistemic formulas – in an infinitary version of the above epistemic language whose technical details are irrelevant to us here – iff there exists a *bisimulation* between ***M*** and ***N*** connecting world *s* to world *t*. We will not pursue the issue whether this is the only possible answer as an account of information structure, but on our view, a true logical understanding of information should definitely come with a 'criterion of identity' allowing us to recognize the same information under different representations.

Logic as information calculus Next, what about the 'calculus of information' provided by epistemic logic? First take the above semantics. Evaluating (model checking) formulas can be done in a systematic way telling us who has which information in a given setting. Likewise, computing updates is a way of calculating information flow. A second major asset are complete axiomatic calculi describing all valid principles of inference with knowledge, such as multi-agent '*K*' or ' *S5*' or other well-known systems. E.g., here is the complete set of principles for *S5*, on top of any complete classical propositional logic:

$K_j(\phi \rightarrow \psi) \rightarrow (K_j\phi \rightarrow K_j\psi)$	*Knowledge Distribution*
$K_j\phi \rightarrow \phi$	*Veridicality*
$K_j\phi \rightarrow K_jK_j\phi$	*Positive Introspection*
$\neg K_j\phi \rightarrow K_j\neg K_j\phi$	*Negative Introspection*

The complete system describes agents' own reasoning, and our reasoning as theorists about them. And here are the required additional axioms for *common knowledge*:

$C_G\,\phi \leftrightarrow \phi \,\&\, E_G\,C_G\,\phi$	*Equilibrium Axiom*
$(\phi \,\&\, C_G\,(\phi \rightarrow E_G\,\phi)) \rightarrow C_G\,\phi$	*Induction Axiom*

Here $E_G\,\phi$ says that everyone in the group *G* knows that ϕ. The complete logic with all these principles is decidable, be it computationally more complex than its propositional base. We will briefly discuss a few of these epistemic axioms below. These laws can be applied in every concrete scenario, just like the principles of probability theory.

Agent Diversity Now there may be a slight unease here, as there is not one calculus, but many. But that may also be counted as a general insight for information theories. Our account of 'knowledge' as 'to the best of an agent's information' really shows that there is an unavoidable interplay between two notions which needs to be made explicit: (a) the information in a situation per se, and (b) the *powers of the agents* having access to it:

> *To see what information is available, one must take*
> *the informational nature of the agents into account.*

Existing epistemic logics differ on the semantic assumptions they make about agents along dimensions (cf. [Liu 2006]) such as deductive powers, introspective abilities, observational powers, and memory capacity. This leads to precise correspondences between special assumptions on information-processing agents and axioms of the logic. Examples are the much-discussed link between the '*KK* principle' $K\phi \rightarrow KK\phi$ and *transitivity* of accessibility, or more recently, that between *commutation* laws for communication (cf. Section 3) and memory assumptions of *perfect recall*. These correspondences provide a refined logical view of what information can flow in a given setting, given the nature of a source and that of the recipient. For an overview of complete epistemic logics over various model classes, see the standard literature.

Alternatives The use of possible worlds models has been under attack from the start. While part of this criticism seems a misguided response to an unfortunate metaphor for the 'situations' represented in a model, there are indeed respectable alternative traditions. Indeed, the earliest semantics for modal logic in the 1930s used *topological models*, reading $K_i\phi$ as 'ϕ is true throughout some open neigbourhood of the current point'. This represents another geometrical intuition about the structure of knowledge as range, generalizing the above graph-based models with accessibility arrows to also include structures like the reals, or Euclidean space. Topological semantics still seems of use ([van Benthem & Sarenac]), and generalized versions show a modest recent flowering in the modern renaissance of 'neighbourhood semantics' ([Arlo-Costa & Pacuit]).

2.3 Other attitudes: doxastic and conditional logic

In the above, we have noted a separation between information per se coming from some source, and agents' attitudes and responses to it. Indeed, there are many attitudes that agents can have toward propositions beyond knowledge. Natural language has a finely-grained repertoire of knowing propositions, believing, or even just entertaining them. Moreover, we can also doubt propositions, maybe on the basis of new information – and more generally, change our allegiance from one epistemic attitude to another. Some of these phenomena have received formal treatment in logic. In particular, in 'doxastic logics', one analyzes assertions $B_i\varphi$ for '*agent i believes that* φ'. The semantics for this operator adds a new idea to the mere information ranges (for different agents) in our epistemic modeling so far. We now assume further gradations, in the form of a *plausibility ordering* of worlds as seen from some vantage point:

$\leq_{i,\,s}\, xy$ in world s, agent i considers y at least as plausible as x.

Thus, while the earlier ranges of epistemic alternatives corresponded to the strict information that we have, the same ranges ordered by plausibility give finer gradations. In particular, we now define belief semantically as *'truth in the most plausible options'*:

$$\boldsymbol{M}, s \models B_i\phi \text{ iff } \boldsymbol{M}, t \models \phi \text{ for all } t \text{ which are maximal in the ordering } \lambda xy.\ \leq_{i,\,s} xy.$$

Here is an elementary example. Consider a model with two worlds that are epistemically accessible, but the one with ¬P considered more plausible than the other:

$$\boldsymbol{P} \xrightarrow[\leq]{\qquad} \neg P$$

In this model, at the actual world where ***P*** holds, the agent does not know whether *P*, but she does (mistakenly!) believe that ¬*P*. It is crucial that our beliefs can be false.

There are some technical complications in making this work in infinite models, which we ignore here. Now, as with epistemic logic, there are complete doxastic logics, and a whole further theory around them [Fagin et al. 1995]. In general the resulting logics also analyze the *interplay* between knowledge and belief, with knowledge implying belief, and more controversially, taking a stance on such issues as whether one knows one's beliefs, and so on. While inter-relations between attitudes toward information are an important topic in logic, we omit it here as tangential to our main concerns.

More interestingly, from a conceptual point of view, we can think of epistemic-doxastic information models with two relations $\sim_i$, $\leq_j$ in two ways, one more *objective* and one more *subjective*. The orderings $\leq_{i,s}$ may just encode an agent's private response to the strict information received, making the belief subjective. But we can also view them more objectively as encoding intrinsic gradations of plausibility in the incoming information – with beliefs the mere registering of this by sensitive logical observers. But from a more technical perspective on information in logic, the following issue may be more pressing. In doxastic logic, one soon finds that mere beliefs are not sufficient for explaining agents' behaviour over time. We want to know what they would do in certain scenarios where they receive new information. This requires *conditional belief*:

$\boldsymbol{M}, s \models B_i(\phi \mid \psi)$ iff $\boldsymbol{M}, t \models \phi$ for all worlds t which
are *maximal* for $\lambda xy.\ \leq_{i,\,s} xy$ *in the set* $\{u \mid \boldsymbol{M}, u \models \psi\}$.

Conditional beliefs $B_i\, (\phi \mid \psi)$ are like logical conditionals in general, in that they express what might happen under different circumstances from where we are now. In particular, they *pre-encode* beliefs in ϕ that we would have if we were to learn new things ψ. The analogy is so close that conditional belief on reflexive transitive plausibility models satisfies exactly the principles of the minimal conditional logic [Burgess, Veltman]. We will return to this setting later on when discussing the mechanics of belief revision.

Hard and soft information Summing up, current epistemic-doxastic-conditional logics paint the following picture. There is *hard information* represented in the current range of uncertainty, i.e., all epistemically accessible worlds. The corresponding informational stance toward this would be the propositional attitude of knowledge, which takes this whole range into account. Next, further fine-structure may be present on these information ranges, through plausibility orderings. These can be viewed either subjectively as representing agents' attitudes, or more objectively, as the result of receiving what might be called *soft information*. This plausibility ordering is reflected in the informational attitude of belief, both absolute and conditional, telling us – roughly – what agents would believe when confronted with new hard information. And in the end, it is the total interplay of all these attitudes which would describe our stances toward information, and the way they are affected by new incoming information. Thus, we get a mixture of information per se, and the ways in which agents take it – and this logical entanglement is such it is hard to say where one notion ends and the other begins.

Remark It has to be said that other views exist, making a much sharper distinction between a base level of pure information processing and a higher level of beliefs on the basis of this information, where belief changes occur in a sort of 'reflective dynamics'. Cf. [chapter Rott] for an architecture of this sort, and [Castelfranchi & Lorini] for an account of triggers like 'surprise' taking us from information update to belief revision.

2.4 Connections with other fields

Epistemology Epistemic-doxastic logic is a somewhat austere mathematical account of qualitative information. But ever since its birth, it has become part of larger discussions in philosophical epistemology. It was soon found that the universal quantifier in information as range provides a rather poor analysis of knowledge in the philosopher's demanding sense, where the quest for a satisfactory definition of *knowing that P* involves finding the right sort of 'robustness' in addition to the obvious features of *truth of P* and *belief in P*. Plato famously proposed 'justified true belief', but sophisticated new definitions have kept appearing until today, such as 'true belief grounded in correct information' ([Dretske]), or 'true belief with counterfactual backing: if *P* had been false, we would have believed that $\neg P$' ([Nozick]). Even though epistemic logic has never offered a definition of knowledge of comparable sophistication, the very mismatch with the richer philosophical tradition has been much more exciting than many happy marriages, leading to clarification and fruitful debate. For instance, the modal distribution axiom $K(\phi \to \psi) \to (K\phi \to K\psi)$ has sparked debates about the idealization of Logical

Omniscience, and the earlier axiom $K\phi \rightarrow KK\phi$ about that of Positive Introspection. And many further points of contact exist. We refer the reader to [Stalnaker], [Lewis], [Hendricks 2005][Williamson 2000], [van Benthem 2006], [Baltag, van Benthem & Smets 2007] for a discussion of a wide array of topics, such as definitions of knowledge, skeptical arguments (see Dretske's chapter in this Handbook), different sources of information (language, senses, etc.), omniscience, bounded rationality, and reflection. Many of them are high-lighted by persistent puzzles, such as Moore sentences, the Fitch Paradox [Salerno], or the Ramsey Test [Ramsey], [Gaerdenfors].

Computer science, and economics But maybe the philosophers set their standards of knowledge too high. Who searches for the Perfect Knight might never get married. For many practical purposes, the above picture of knowledge as range seems quite sufficient. Think of scenarios like this. The cards have been dealt. I know there are 52 of them, and I know their colours. The possible worlds are just the possible deals. Of course, I could be wrong about this (perhaps someone replaced the King of Hearts by Bill Clinton's visiting card), but this worry seems morbid, and not very useful in understanding normal information flow. What is true is that less well-founded attitudes will come in as well. I may only have ephemeral beliefs about who holds which card, or about how the other agents will play. And indeed, we are sensitive to such distinctions. 'Knowledge' may then be the right term for 'the strictest attitude in the current setting'.

Indeed, notions of knowledge based on epistemic logic have penetrated other areas, notably computer science, witness the TARK conferences since the early 1980s [Fagin et al. 1995]. Originally, knowledge was ascribed here to processors in distributed systems, with accessibility arising since these can only distinguish global system states through their own local state. But in modern agent-based computing systems, the difference with humans seems slight. Also, epistemic logic entered game theory in the 1070s ([Aumann]), as a way of stating what players know about each other, in an account of the reasoning about rationality underpinning for notions like Backward Induction and Nash Equilibrium. (Cf. [Stalnaker], [Halpern], [van der Hoek & Pauly].)

In applications such as these, the account of knowledge often occurs intermingled with accounts of moves or actions. Players reason using their knowledge of what certain moves will bring about, and also, after observing a move by other players, they readjust their current information. This natural combination will be the topic of the next section.

3 Information flow and dynamic logic

Structure and process As we said before, it is hard to think of information in isolation from the processes which create, modify, and convey it. This combination of structure and process is natural in many disciplines. In computer science, one designs data structures in tandem with the processes that manipulate them, and the tasks which the latter should perform. But the same point is familiar from philosophy ([David Lewis]), when saying that 'Meaning Is what Meaning Does'. We can only give good representations of meanings for linguistic expressions when we state how they are going to be *used*: in communication, disambiguation, inference, and so on. In a slogan:

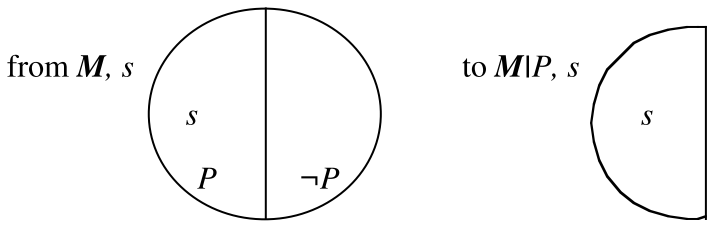
> *Structure should always come in tandem with a process!*

Which processes drive the notion of information? There is a great variety of candidates, witness the various chapters of this Handbook: computation, inference, update, revision and self-correction, question answering, communication, interaction, games, learning, and so on. Some of these processes concern activities of single agents, whereas others are intrinsically multi-agent 'social' phenomena – with interesting connections between the two, as in the passage from argumentation to formal proof. We will not investigate all these structure-dynamics links in detail, or even the issue whether one notion of information serves all of them [cf. the Editorial to this Handbook].

Instead, we will highlight one instance of this Dynamic Turn, in terms of informational processes for epistemic-doxastic logic (see [Baltag et al., Rott] for further elaboration).

Information update and model change One of the simplest events providing new information is a *public announcement !P* of some fact *P*. One can think of this as a real statement coming from some authoritative source – or more generally, as a totally reliable observation based on one's senses. If I see that the Ace of Spades is played on the table, I now come to know that no one of us holds it any more. Of course, we can be mistaken about what we hear or see, but morbid worries can block the progress of logic. These events of new *hard information* in the earlier sense change what I know. Formally, this happens by their triggering changes in the current epistemic model.

In particular, public announcements *!P* work as described in Section 1 above. They eliminate all worlds incompatible with *P*, thereby zooming in on the actual situation. Thus the current model $(\boldsymbol{M}, s)$ with actual world s changes into its submodel $(\boldsymbol{M}|P, s)$, whose domain is the set $\{t \in \boldsymbol{M} \mid \boldsymbol{M}, t \models P\}$. In a picture, one goes



Crucially, truth values of formulas may change in such an update step: agents who did not know that P now do after the announcement. One can keep track of this in *public announcement logic PAL*, extending the epistemic language with action expressions:

Formulas	P:	$p \mid \neg\phi \mid \phi\vee\psi \mid K_i\phi \mid C_G\phi \mid [A]\phi$
Action expressions	A:	$!P$

The semantic clause for the dynamic action modality is as follows:

$$\boldsymbol{M}, s \models [!P]\,\phi \quad \text{iff} \quad \text{if } \boldsymbol{M}, s \models P, \text{ then } \boldsymbol{M}|P, s \models \phi$$

As all this is less known than standard epistemic logic, here is the complete calculus of information flow under public announcement. It has the usual laws of epistemic logic over one's chosen base models plus, crucially, the following *reduction axioms*:

(a)	$[P!]q$	$\leftrightarrow$	$P \rightarrow q$	*for atomic facts* q
(b)	$[P!]\neg\phi$	$\leftrightarrow$	$P \rightarrow \neg[P!]\phi$	
(c)	$[P!]\phi\wedge\psi$	$\leftrightarrow$	$[P!]\phi \wedge [P!]\psi$	
(d)	$[P!]K_i\phi$	$\leftrightarrow$	$P \rightarrow K_i(P \rightarrow [P!]\phi)$	

Logical axioms high-light basic issues in concentrated form. In particular, the PAL reduction axioms allow for stepwise compositional analysis of the effects of assertions. Here, the key axiom (d) relates the knowledge agents have following new information to knowledge they had before. This highlights two important issues. The first is the earlier idea of *pre-encoding*. There is already sufficient information about the effects of events $!P$ in the current state through the *relativized knowledge* of the form $K_i(P \rightarrow$.

But also, axiom (d) has a remarkable *interchange* between two operators: the dynamic $[!P]$ and the epistemic K_i. This commutativity is not obvious, witness general logics of knowledge and action (Moore 1985). I know now that after drinking I am terribly boring. But the tragedy of my life is that, after drinking, I do not know that I am boring, and I go on and on and on... The reason is that drinking, unlike mere observation, impairs my epistemic abilities. Knowledge action interchanges presuppose abilities of agents such as Perfect Recall ([Halpern & Vardi 1988], [van Benthem 2001]). Again we see that studying information flow immediately raises issues of agent ability.

As a logic, *PAL* has a modal bisimulation-based model theory, and it raises many new issues of expressive power and computational complexity (cf. van Benthem 2005).

General observation and event update But public announcement or public observation is just one process which produces information. More complex changes occur in the current information model when agents have only partial powers of observation. Say, I see you draw a card from the stack, but unlike you, without seeing exactly *which one*. In general *dynamic-epistemic logics* (cf. [Baltag, Moss & Solecki, Gerbrandy, van Benthem, van Eijck & Kooi]), *event models* ***A*** can describe complex scenarios where not all agents have the same observational access to what is happening. This leads to a mechanism of *product update*, turning the current epistemic model ***M*** into a model ***MxA*** which can even be larger than ***M*** itself, recording information of different agents about facts and what others know. In particular, product update redefines the universe of relevant possible worlds, and also the epistemic accessibility relations between them. Conversations, games, internet transactions, and other real activities are like this. For this natural continuation of our present analysis, we refer to the chapter [Baltag et al.].

Remark Systems with similar aims are *epistemic-temporal* logics, proposed by Gupta-Thomason, Belnap et al., Halpern et al., Parikh & Ramanujam, etc. [van Benthem & Pacuit 2006], [van Benthem, Gerbrandy & Pacuit 2007] give surveys and comparisons.

Changing beliefs The dynamic process perspective on information change explained here also applies to the doxastic attitude of our *beliefs*, and how to *revise* these on the basis of incoming information ([Gaerdenfors 1987]). This involves changes, not in the range of available worlds or in their epistemic accessibility patterns, but rather in the earlier *plausibility orderings* $\leq_{i,\ s} xy$ among worlds. How this works precisely, depends on the incoming signal. When we receive hard information *!P*, update will proceed by world elimination as before. We then get new beliefs related to our earlier conditional beliefs, and the counterpart to the above reduction axiom (d) are laws saying which new beliefs – and indeed, conditional beliefs – are acquired ([van Benthem 2006]):

(e) $[!P]\, B_i\phi \quad \leftrightarrow \quad P \rightarrow B_i([!P]\phi \mid P)$

(f) $[!P]\, B_i(\phi \mid \psi) \quad \leftrightarrow \quad P \rightarrow B_i([!P]\phi \mid P \wedge [!P]\psi)$

But often, we just get *soft information*, of the sort mentioned in Section 2. This just increases our 'preference' for *P*-worlds, without totally abandoning the others. A quite typical 'belief revision policy' in this spirit is *lexicographic upgrade* $\Uparrow P$ (Rott 2005) which replaces the current ordering relation $\leq$ between worlds by the following: *all P-worlds become better than all ¬P-worlds, and within those two zones, the old ordering remains*. Belief changes under such policies can again be axiomatized completely ([van Benthem 2007]). Just for illustration, here is the crucial axiom for the widely used $\Uparrow P$ revision policy. It looks somewhat forbidding – but perhaps necessarily so: after all, we are now describing a more complex informational process than mere epistemic update:

(g) $[\Uparrow P]\, B(\phi \mid \psi) \ \leftrightarrow\ (\langle\rangle(P \wedge [\Uparrow P]\psi) \wedge B([\Uparrow P]\phi \mid (P \wedge [\Uparrow P]\psi))$
$\vee (\neg(\langle\rangle(P \wedge [\Uparrow P]\psi) \wedge B([\Uparrow P]\phi \mid [\Uparrow P]\psi))$

Here <> is the existential epistemic modality saying 'in some world of the current range'. Still richer dynamic doxastic logics use the above event models, with 'signal events' ordered by plausibility (cf. [Baltag & Smets 2006]). This allows for relocating revision policies in the structure of the incoming signal, while sticking to one product update for the new plausibility ordering. We refer to the chapters [Baltag et al.], [Rott] in this Handbook for further revision policies, and deeper discussion of the issues raised by all this. Finally, there are also strong analogies between plausibility reordering and recent dynamic logics of *preference change* ([van Benthem & Liu 2006]).

The general setting In addition to the mentioned chapters on information update and belief revision, the above account of epistemic and doxastic processes is in line with broader current trends. These include 'logical dynamics' ([van Benthem 1996]), 'information dynamics' in computer science (our chapter [Abramsky]) and 'interactive epistemology' in philosophy and game theory (our chapter [Walliser]). There are also obvious connections with mathematical *learning theory* (our chapter [Kelly]). As a general conclusion we note once more that the resulting logics of information have the following features as 'information theories'. They are *dynamic*, in that they represent informational actions and events on a par with static information structures. They are also *social*, in that most basic scenarios involve more than one agent together. Even observation is really a matter for two agents: me and my source. And notions of group knowledge are crucial to understanding information flow. This dynamic social character of information comes together in the central concept of *interaction* between agents.

4 Information as Correlation: the world of Situation Theory

We now turn to the only program in modern logic that treats information as its central notion, shaking up the usual agenda of research. Situation theory is based on how the structure of the world underlies the existence and flow of information. It originated in the 1980's [Barwise & Perry] as a foundational framework for 'situation semantics', but by now the program has come to study the notion of information from a much more general point of view (cf. also [chapter Devlin & Rozenberg]).

Situation theory goes beyond possible worlds models by providing richer structure. In particular, this yields a natural treatment of the problem of grain size, with different *false* claims about the world represented by different mathematical objects. Other important concerns that pushed the theory were context effects in inference and language. According to situation semantics, meaning arises from the interaction of organisms and their information-rich environment. This reflects the influence of Ecological Realism ([Gibson]), where visual perception involves the world as a structured reality full of (in situation-theoretic terminology) 'uniformities' to which organisms are 'attuned' in order to survive. Thus, information is a pervasive aspect of reality, prior to cognitive action.

Other important influences along the same line came from philosophy. For example, Putnam's Twin Earth thought experiment ([Putnam]) has convinced many that meaning cannot be just in the mind, and hence external reality must play a crucial role. But it is Dretske's theory of information flow (cf. [Dretske]) what more conspicuously provided key ideas about the basic informational notion to study. Dretske builds on Shannon's theory of information transmission, but his account is qualitative and his focus is squarely on semantic *content*, not on the *amount* of information transmitted. It uses the following variant of the usual notion of confirmation:

> **Information Content** To an agent with prior knowledge *k*, signal *r* carries the information that *s* is *F* iff $Pr(s \text{ is } F \mid r \text{ and } k) = 1$, but $Pr(s \text{ is } F \mid k) < 1$.

Dretske then defines knowledge as belief caused by information flow. One reason why the definition includes the strict equality $Pr(s \text{ is } F \mid r \text{ and } k) = 1$, which has been the target of some criticisms, is the following intuitively appealing consequence:

> **Xerox Principle** If *A* carries the information that *B* and *B* carries the information that *C*, then *A* carries the information that *C*.

Situation theory adopts the same guiding idea – though dropping the probabilities –, and it encompasses not only natural but also conventional signals, among others.

We now come to the basic apparatus, which we develop in quite some detail, also for its interesting contrast with the preceding sections. Epistemic logic in its static and dynamic versions has a well-developed model theory and proof calculus, but it provides little information at a practical level about what might be called the 'art of modeling'. Situation theory does a kind of converse: it offers a rich apparatus for modeling challenging phenomena, though it has (with a few exceptions to be noted below) not yet generated canonical calculi of reasoning of the same sweep and wide application as epistemic logic.

4.1 Basic concepts of situation theory

Depending on long-term adaptive mechanisms and short-term goals, agents carve out reality in many different ways. Their schemes of individuation allow them to recognize uniformities of the world, including *individuals, relations*, spatial and temporal *locations*, and in particular, *situations*: parts of reality that can be singled out as such by agents as relevant to some enquiry, or some task at hand. Those parts of the world are of certain *types*: for individuals, locations, etc. As for predication, each relation has some *roles* associated with it. For example, the relation "kisses" has two roles: (1) the person who kisses and (2) the one kissed. Only individuals can fill these two roles.

Basic infons Situations can be classified via *infons*, i.e., 'pieces of information'. The simplest infons are called *issues*, and they consist of a relation, an assignment of *roles*, and a *polarity* (0 or 1, intuitively standing for 'false' and 'true'). For example,

♦ = << *kisses, Anne, Bill ; 1* >>

is an issue. As for the relation between information and situations, we say that

s |= ♦ (read as *"s supports ♦"*)

if ♦ is made factual by *s*. With ♦ as above, this happens exactly if *s* is a situation where Anne kisses Bill. A situation *s* is said to resolve an issue ♦ if it supports either the issue or its dual (the issue that is just like ♦ but with the opposite polarity). As situations are only parts of reality, they will in general fail to resolve all issues.

Compound infons The next obvious aspect of structure are ways of constructing infons. A guide for assessing these is the Principle of Persistence, based on the natural ordering s_1 🗁 s_2 among situations of "being a part of". Persistence says that all information supported by a situation s_1 must be supported by any more inclusive situation s_2. That is, if s_1 🗁 s_2 and s_1 |= ♦, then s_2 |= ♦. In particular, given infons ♦ and ◆, we are justified to create a compound infon ♦ ⮝ ◆ with *s* |= ♦ ⮝ ◆ iff *s* |= ♦ and s |= ◆. This is so because the resulting infon is persistent. The same holds for disjunctions and other constructions. (This is like the semantics of intuitionistic logic, and partial logics: cf. the corresponding chapters in the Handbook of Philosophical Logic.)

Of course, one would like an algebra and logic of infons. We will not go in details of that (cf. [Devlin]), but just point out that the usual logical intuitions cannot be taken for granted here. For an illustration (cf. [Moss & Seligman]), suppose we say that

♦ ⮝ ◆= ♦' ⮝ ◆' if (♦ = ♦' and ◆ = ◆') or (♦ = ◆' and ◆ = ♦').

Combining this with our logical intuitions about conjunctions, we get by idempotence of ⮝, ♦⮝◆ = (♦⮝◆) ⮝ (♦⮝◆), and by our equality condition, ♦ = ♦ ⮝ ◆. This mismatch with classical intuition reflects the fact that *information is fine grained*: logically equivalent compound infons may still be different pieces of information. Thus, a notion of equality of compound infons cannot be totally based on syntactic properties.

Abstract infons Before we talk about information flow, we need two more notions: *parameters* and *situation types*. Parameters are constructs resembling the well-known operator of ●-abstraction in the ●-calculus ([Barendregt]). We start by adding some infinite set of parameters for each basic type. Say we add parameters $â_1$, $â_2$, ① of the type of individuals, _$_1$, _$_2$, ① of type of relations, etc. These parameters are like indeterminates that can take as values any individual, relation, etc., as the case may be.

Now, *abstract infons* are just like infons, but some roles may be filled with parameters instead of concrete objects. Abstract infons may be seen as properties. For example,

♦ = << *kisses,* $\hat{a}_1$, $\hat{a}_2$; 1 >>

captures the property shared by all situations where somebody is kissing somebody else. In this setting, a new information-related ordering beyond inclusion of situations emerges. Parametric infons are naturally ordered by *specificity*. E.g., the abstract infon

♦' = << *kisses, Anne,* $\hat{a}_2$; 1 >>

is more specific than ♦. This notion of specificity is made formal via 'anchors': partial functions mapping parameters to objects of the appropriate types. For an abstract infon ♦ and anchor *f*, ♦[*f*] is the parametric infon obtained from ♦ by replacing the parameters that appear in both ♦ and the domain of *f* by their *f*-values. E.g., with ♦ and ♦' as in our example, we have ♦' = ♦[$\hat{a}_1$ ✸ *Anne*]; so ♦' is more specific than ♦.

One can extend the earlier support relation |= to parametric infons. If *s* is a situation and ♦ an infon (parametric or not), we define *s* |= ♦ iff there is some anchor *f* such that ♦[f] is an infon and *s* |= ♦[*f*]. If *s* is a situation where Anne is kissing Bill, then

s |= << *kisses, Anne,* $\hat{a}_2$; 1 >>

because *s* |= << *kisses, Anne,* $\hat{a}_2$; 1 >> [$\hat{a}_2$ ✸ *Bill*], so

s |= << *kisses, Anne, Bill*; 1 >>.

Situation types Once we have parametric infons, we can abstract over situation parameters, and thus obtain a rich collection of *abstract situation types*, of the form

[_ | ♦[*f*] is an infon for some anchor *f* and _ |= ♦[*f*]],

where _ is a situation parameter and ♦ is an abstract infon. The extension of this type is given by all situations that match ♦: that is, the set

{ *s* | *s* is a situation and for some anchor *f*, ♦[*f*] is an infon with _ |= ♦[*f*] }

Viewed conversely, the type is a property shared by all situations in its denotation. If *T* is a parametric type, we use *T*(***p***) to stress the fact that *T* came from abstracting over a parametric infon ♦ with parameters ***p***. If *f* is an anchor, then *T*[*f*] is the type that results from abstracting over the parametric infon ♦[*f*]. Besides abstracting from parametric types, there are several other ways of constructing new types (cf. [Devlin]).

Constraints So far we have an ontology of the world as a reality which can be grasped in its parts, the situations. Here situations are not 'modal' alternative possible worlds: they are structured parts of one world, they support pieces of information, and they are of certain types. What is still missing in the picture, however, is the following crucial fact (cf. [Dretske], [Barwise & Perry]). Reality is full of lawlike *constraints* between types of situations, which make it possible for one situation to carry information about another. A *constraint* is simply a *relation between (*parametric or non-parametric) *types* that can in principle be formalized as the simple infon

$$<< Involves, T\,[\boldsymbol{p}], T'; 1 >>.$$

This infon is factual if the following existence condition is satisfied:

for every anchor f with domain $\boldsymbol{p}$, if s is a situation of type $T[f]$,
then there is a situation s' of type $T'[f]$.

So if *SMOKE* is the type of those situations where smoke is present, and *FIRE* is the type of those situations where something is burning, then << *Involves*, *SMOKE*, *FIRE*; 1 >> is factual. Constraints correspond to natural phenomena such as natural laws of physics, social conventions in language and traffic signs, and any other sort of dependence.

As a point of realism, most regularities to which agents are attuned hold only within some regions of reality. If a massive object is released in the air, it will fall, *given that* we are on Earth, not in a satellite, say. *Conditional constraints* are needed to deal with this. They can be seen then as ternary relations on types, formalized by infons

$$<< Involves, T_1\,[\boldsymbol{p}], T_2, T_3\,[\boldsymbol{q}] >>.$$

Such a constraint is factual if for every anchor f with domain $\boldsymbol{p} \cup \boldsymbol{q}$, if s is a situation of types $T_1[f]$ and $T_3[f]$, then there is a situation s' of type $T_2[f]$. Conditional constraints will also allow us to deal properly with issues of reliability and error in Section 5.4.

4.2 Information flow, regularities, and types

Now, let us get to the fundamental issue of this Section: *how can one situation carry information about another*, which may be far away in time or space? We present two ways of formalizing this claim. The first is shorter, in terms of the above ontology. The second requires the theory of classifications and channels of [Barwise and Seligman].

Propositions and types In Situation theory, the bearers of truth are *propositions* – and there are two kinds of these. *Russelian* propositions are characterized just by a type T, and claim that there exists a situation of type T. *Austinian* propositions are claims of the form $s \models T$, involving a type plus a situation. Now, while propositions are the bearers of truth, it is particular situations that act as carriers of information. More precisely, it is the fact that some proposition $s \models T$ is true, plus the existence of factual constraints relating T with other types, which allows for s to carry information about other parts of reality. The basic kind of informational statement for this reads:

The fact that *s* |= *T* carries the information that *Prop.*

Whether the proposition *Prop* is Russelian or Austinian tells us whether the information carried by *s* is what may be called 'pure' or 'incremental'. We say that

s |= *T* carries the pure information that *T'* if there
is a factual constraint << *Involves T, T';1* >>.

For example, starting with the Russelian case, the fact that Anne is kissing Bill carries the pure information that she is touching somebody, because the constraint

<< *Involves*, *kisses*[x, y], *touches*[x, z] ;1 >>

is factual. Next, Austinian propositions characterize incremental information. We say

s |= *T* carries the incremental information that *s'* |= *T' (relative to T")*
if there exits a factual conditional constraint << *Involves, T, T', T"; 1* >>.

Thus, that Anne is kissing Bill carries the incremental information that she is touching Bill, since the constraint << *Involves*, *kisses*[x, y], *touches*[x, z], *sameperson*[y, z] ;1 >> is factual. The type *T"* in the conditional constraint describes how the carrier situation and the described situation *s*' are connected. Incremental information therefore, is more specific than pure information. It gives information about a *concrete* situation *s'* via the indicating fact *s* |= *T*, in virtue of how *s* is connected to *s'*. [Israel and Perry] argue for the utility of both incremental and pure information to guide behavior. E.g., the city-alarm being triggered at a non-rehearsal time carries the pure information that there is *some* threat for the citizens, information that might save lives.

Summarizing, here are three fundamental principles about the nature of information. We state them in terms of 'distributed systems', rather than reality, because essentially, these principles apply to any kind of system that can be analyzed in sub-parts:

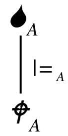
(1) The availability of information depends on the existence of regularities between connected parts of a distributed system.
(2) These regularities are relative to how a distributed system is analyzed in parts.
(3) Flow of information crucially involves types and their concrete instances. It is in virtue of constraints (relations on types) that concrete situations, being of certain type, can act as carriers of (concrete or general) information.

Agents, channels, and information flow once more There is an ingredient missing in the account so far, which becomes apparent when we think of information as used by *agents* to guide their actions. Then it is clear that information is also relative to the manner in which an agent is adapted to the world. This means that the information an agent can get out of a situation depends to the constraints to which it is *attuned*.

4.3 Classifications and channel theory

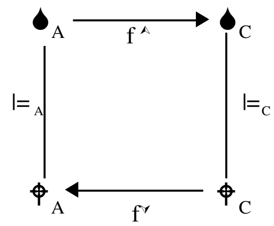
Several mathematical models of situation theory have been developed, based on axiomatics or (non-well-founded) set theory. But to explore the handling of information by agents in more depth, we will now sketch the theory of classifications and channels of [Barwise & Seligman], whose formal foundation is category theory.

Classifications (independently discovered around 1990 as [Chu Spaces] and [Concept Lattices]) are the basic notion of the theory. They are triples, often depicted as follows:



Here ⌖$_A$ is a set of *tokens* (for example situations), 🌢$_A$ is a set of *types* (conceived of as anything that can classify tokens), and |= is a relation between tokens and types. If *s* is a token and *T* a type, then $s \models_A T$ reads '*s* is of type *T*', just as before.

While classifications model components of a system with their local properties, we must also model their natural 'part-of' and 'flow' relationships. Here, the basic structural invariance are *infomorphisms.* If ***A*** = < ⌖$_{A,}$ 🌢$_{A,}$ $\models_A$> and ***C*** = < ⌖$_{C,}$ 🌢$_{C,}$ $\models_C$ > are classifications, an infomorphism *f: **A*** ✸ ***C*** is a pair $< f^{\wedge}, f^{\vee} >$ of functions



such that for all tokens *b* ♏ ⌖$_C$ and all types *T* ♏ 🌢$_A$

(*) $f^{\vee}(c) \models_A T$ if and only if $c \models_C f^{\wedge}(T)$

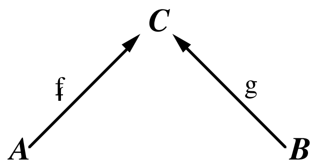
Infomorphisms are of independent interest as the abstract invariance behind translation between theories, and that of general category-theoretic adjunctions. The structural properties of classifications preserved by them were studied in [van Benthem 2000]. But here we look at their concrete uses as an intuitive model for information flow.

A concrete scenario Modeling the smoke-in-the-mountaintop scenarios of the Indian logic tradition, mentioned in Section 1, involve at least two classifications **A** and **C**. The tokens in ⌖$_A$ might be situations where somebody is facing a mountain, with the 🌢$_A$ types such as SEESSMOKE, LOOKINGUP, LOOKINGDOWN, BLIND etc. On the other hand, the classification **C** might be the overall setting including the observer and the mountain. Its tokens are situations that extend those of ⌖$_A$, and types in 🌢$_C$ might be things such as

OBSERVERSEESSMOKE, THEREISAFIREONTOP, etc. The map $f^{\vee}$ maps each large situation to the sub-situation capturing just the point of view of the observer. The map $f^{\wedge}$ sends SMOKEOBSERVED to OBSERVERSEESSMOKE, LOOKINGDOWN to OBSERVERLOOKINGDOWN, etc. Thus type *T* of **A** is mapped to a type of **C** intended to mean that 'the observing situation' is of type *T*. Condition (*) ensures that things work this way.

Regularities between types that hold in a classification are like our earlier 'constraints. Formally, we redefine a *constraint* $♠_1$ ○ $♠_2$ *of classification* **A** to consist of two sets of types such that for all *a* ♏ $⚲_A$, if *a* $|=_A$ ∧$♠_1$, then *a* $|=_A$ ∨$♠_2$. If **A** is still our classification of observers facing a mountain, then SEESSMOKE, BLIND ○ ∅ would be a constraint of **A** that says that no blind observer sees smoke on top of the mountain.

Channels and information flow Let us now add to our observer and his mountain a third classification ***B*** for what is happening at the mountain top. Its tokens are situations located on mountain tops, and its types include things such as THEREISFIRE, THEREISFOG, etc. ***B*** is also a 'part' of the big component ***C*** say via the infomorfism *g:*



A collection of infomorfisms sharing codomain *C* is called a *channel* ♣ with *core C*. Tokens of the core are called *connections*, because they connect subparts into a whole. Tokens *a* from $⚲_A$ and *b* from $⚲_B$ are *connected* in channel ♣ if there is a token *c* ♏ $⚲_C$ such that $f^{\vee}(c) = a$ and $g^{\wedge}(c) = b$. In the example, an observing situation and a mountain top would be connected precisely if they belong to the same overall situation. We can now formulate a notion of (incremental!) information flow:

> *a* $|=_A$ *T carries the information that b* $|=_B$ *T '* (relative to ♣) if
> *a, b* are connected in ♣ and $f^{\wedge}(T)$ ○ $g^{\wedge}(T')$ is a constraint of ***C***.

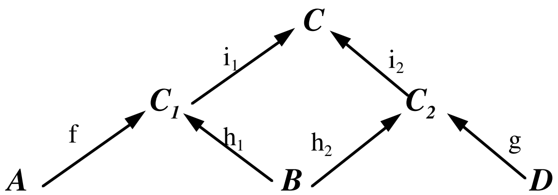
This notion of information flow is relative to a channel – and hence, to an analysis of a whole into parts. Again we see that “carrying information" is not an absolute property: the mere fact that a token or situation is of a certain type does not completely determine what information it carries. This feature links up with the study of context dependency in logic (cf. Section 5).

Here is our observer-mountain example in these terms. Let *s*♏A be a situation of type SEESSMOKE: the observer in it sees smoke on top of the mountain. Let *s'*♏ B be the top-of-the mountain situation observed from the base of the mountain . Then our choice of **C** makes *s* and *s'* linked by some connection in **C**. In addition, $f^{\wedge}$(SEESSMOKE) =

OBSERVERSEESSMOKE and g^(THEREISFIRE) = THEREISAFIREONTOP. Hence, since OBSERVERSEESSMOKE **O** THEREISAFIREONTOP is a constraint of **C**, we have that *s* being of type SEESSMOKE carries the information that *s*' is of type THEREISFIRE.

Xerox Principle revisited Suppose now that we have two channels with cores $\boldsymbol{C_1}$, $\boldsymbol{C_2}$.

We can form their direct sum $\boldsymbol{C} = \boldsymbol{C_1} + \boldsymbol{C_2}$, whose tokens are ordered pairs of tokens of $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$ with types as the disjoint union of those of $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$:



Here the type component of infomorphisms i_1 and i_2 is the identity, and tokens go via left and right projections. We get then this version of the Xerox Principle: if a |=$_A$ T carries the information that b |=$_B$ T' (relative to ♣$_1$) and b |=$_B$ T' carries the information that d |=$_D$ T'' (relative to ♣$_2$), then a |=$_A$ T carries the information that d |=$_D$ T'' relative to the channel with core $\boldsymbol{C}$ and infomorphisms $i_1 f$: $\boldsymbol{A}$ ✹ $\boldsymbol{C}$ and $i_2 g$: $\boldsymbol{D}$ ✹ $\boldsymbol{C}$.

Digression: model-theoretic relations between situations So far, we just *assumed* that a given system has dependencies. But *why* does one situation correlate with another? Sometimes, this is a mere accident, just as funds in the stock market may have fluke correlations. But it is the more intimate links that we are after. A watch keeps time since its stepwise operation mirrors the unfolding of time. The strongest notions that support such harmony are model-theoretic. Infomorphisms are one elegant and general example, and following Barwise & Seligman 1995, van Benthem 2000 has studied them model-theoretically, determining the precise first-order properties of structures which they preserve. But other model-theoretic relations between situations make sense, too, such as *model extension* to larger situations. For a study of model-crossing notions of inference and *logics of model change*, cf. [van Benthem 2007]. Barwise & van Benthem 1999 study general 'information links', and entailment along model-crossing relations.

Now, let's turn to explicit rules for reasoning with this semantic apparatus.

Reasoning with local constraints A *local constraint* of $\boldsymbol{A}$ = < ✤$_A$, ♦$_A$, |=$_A$ > on a set of tokens S has the form ✐ |-$_A$ ♦ (on S), where ✐, ♦ are two sets of types such that for all *a* ∈ S, if *a* |=$_A$ ⮝✐, then *a* |=$_A$ ⮟🌢. This generalizes the notion of a conditional constraint. This setting validates the following calculus:

- *Identity*: $\overline{}$ *T* |-$_A$ *T* (on S)

- *Weakening*:

  🌢$_1$ |-$_A$ 🌢$_2$ (on S)
  ____
  🌢$_1$, ✐$_1$ |-$_A$ 🌢$_2$, ✐$_2$ (on S)

- *Global Cut (✐)*:

  🌢$_1$, ✐$_1$ |-$_A$ 🌢$_2$, ✐$_2$ (on S) for all ✐$_1$, ✐$_2$ with ✐$_1$ ⭿ ✐$_2$ = ✐
  ____
  🌢$_1$ |-$_A$ - 🌢$_2$ (on S)

Thus, we can model reasoning by agents who are attuned to only *some* local regularities of a system and its parts, and we can exclude from the logic spurious 'regularities' that may exist out of chance. But to make this more realistic, we would also like to have rules for shifting classifications (changing from |-$_A$ to |-$_B$, say) and conditions (from *S* to some other S'). Some examples from [Martinez] look like this:

Enlargement:

🌢$_1$, T |-$_A$ 🌢$_2$ (on *Nec(T, S)*)
____
🌢$_1$, T |-$_A$ 🌢$_2$ (on S)

where *Nec* is a function that maps a type and a set of tokens *S* to a *subset* of *S* that includes all tokens of type *T* in *S*. Other useful rules include:

- *S-Weakening*:

  🌢$_1$ |-$_A$ 🌢$_2$ (on S)
  ____
  🌢$_1$, ✐$_1$ |-$_A$ 🌢$_2$, ✐$_2$ (on S') whenever S' ⊄ S.

- *Divide and Conquer:*

  🌢$_1$ |-$_A$ 🌢$_2$ (on S) 🌢$_1$ |-$_A$ 🌢$_2$ (on S')
  ____
  🌢$_1$ |-$_A$ 🌢$_2$ (on S ⭿ S')

And these final rules ([Barwise & Seligman]) hold for each infomorphism *f:* ***A*** ✹ ***C*** :

- *f-Intro*:

  ✐ |-$_A$ 🌢 (on S)
  ____
  f [✐] |-$_C$ *f* [🌢] (on (f ⮟) $^{-1}$[S])

- *f- lim*:

  ✐ |-$_C$ 🌢 (on S)
  ____
  f^{-1} [✐] |-$_A$ f^{-1} [🌢] (on f ⮟ [S])

Here *f* [✐] is the set of types obtained by applying *f* ^ to all types in ✐, f^{-1} [✐] is the set of types whose image under *f* ^ is in ✐.

While this inference system is powerful, it still lacks a good account of *inferential information* (cf. Section 6 below) on which [Barwise & Seligman] have only this to say: to an agent with prior knowledge *k*, *r* being F carries the information that *s* is *G* iff the agent can legitimately infer that *s* is *G* from *r* being *F* and *k*, but not from *k* alone.

4.5 Circularity: hyper-sets and reasoning

Situation Theory has been applied to model not just straight information flow, but also self-referential phenomena. For example, the study of natural language as a medium for information conveyance involves situations in which utterances is made. But utterances may be precisely about the situation in which they are made! Now, in set-theoretic approaches to Situation Theory, situations are usually modeled as a set of tuples (standing for the infons supported by the situation), so this kind of self-referring situations cannot be naturally modeled as sets in the standard Zermelo universe of sets. This is so because in the Axiom of Foundation bans the existence of chains of sets of the form $u \in u_1 \in \ldots \in u_{n-1} \in u$. Due to this, it became more and more apparent, that Situation Theory required a theory of sets without Foundation. In fact, an important contribution of Situation Theory has been the study of circularity and its applications.

Technically, most of this work has taken place in the theory of sets *AFA* whose axioms include all those of ZFC (except Foundation), plus the Antifoundation Axiom (cf. [Forty & Honsell], [Aczel], [Barwise & Moss]):

> *Given any graph (W, R) whose childless nodes are labeled with urelements, there is a unique way to label the other nodes with sets such that, if d(v) is the label of v, then d(v) ={ d(u) | v Ru} (i.e., d(v) is the set of labels of v's children).*

Each graph whose childless nodes are labeled with urelements represents a set, and each set can be represented as a labeled graph. In fact, the same set may be represented by different, but *bisimilar* labeled graphs. For example, there is a unique set such that $\Omega = \{\Omega\}$, and it can be represented by the following two graphs (among many others):

Bisimilarity enters the picture because in the AFA universe, equality of elements is not enough as a criterion for deciding whether two sets are equal. This fact leads to interesting connections with modal logic, and eventually to the use of recent *co-algebraic* methods. For instance, AFA has the following existence property ([Aczel]):

Solution Lemma Each system of equations in sets has a unique solution.

Thus, there are unique sets *U, V* which satisfy (with *a, b* urelements) the equations:

$$U = \{U, V, a\} \quad , \quad V = \{V, b\}$$

More abstractly, one can talk about such systems of equations in terms of *co-algebras* and category-theoretic morphisms. For each monotone operator *F* on sets, any mapping $f: X \rightarrow F(X)$ is a coalgebra. This is true, for example, of $F(X) = P(X \cup A)$, an operator directly related to the setting up of simple systems of equations as the one shown above.

A main theme in this direction is the duality between algebras, inductive definitions and smallest fixed points on one hand, and coalgebras, coinductive definitions and largest fixed points on the other hand ([Jacobs & Rutten], [Venema]).

This mathematical apparatus models the above phenomena in a comprehensive manner ([references]). See [Barwise & Moss] for a general theory, involving a mixture of AFA set theory and early co-algebra, and [Baltag] for a study of classes in this framework. A theory of AFA-classes was important as a development, beyond purely technical reasons, because although the initial intuition about situations is that they are 'small', situation theory actually deals with very large set-theoretic objects. For instance, the treatment of paradoxes via Austinian propositions allows for models of reality that resolve all the possible issues, and which turn out to be proper classes [Barwise and Etchemendy].

Other applications Summarizing, situation theory is a logical theory of information in distributed systems (including reality as parsed by agents) that afford phenomena of information as correlation. Again, it brings along its own agenda, which now includes issues of circularity in modeling and the associated reasoning apparatus, and a general account of relations between situations and of the channels along which information can flow. For further applications, we refer to the more linguistic-social [chapter Devlin & Rozenberg]), as well as the computational approaches of [Akman], [Martinez]). Finally, [Devlin] is a good introduction to the general ideas and state of the art in the early 1990s.

5 Merging Range and Constraint Views

So far, we have told two separate stories of information as range and information as correlation. In this section, we will compare and eventually even merge the two. While we cannot do justice to all aspects here – this would require a longer separate publication –, we hope to show that the two styles of thinking are really very congenial. But can we?

Indeed, many readers might rather have expected a nice polemical shoot-out here. It is often thought that epistemic logic and situation theory are hostile paradigms (cf. [Perry vs. Stalnaker]). For instance, much has been made of the *partiality* of situations versus the 'totality' of possible worlds. But in practice, many of these differences are untenable. In many models of epistemic logic, possible worlds are small and partial: witness the card examples in Section 2, where 'worlds' are just possible hands, and hence 'no big deal'. Conversely, some entities cited as 'situations' can be pretty heavy, such as countries, or wars. Also, situation theory has been cast as 'realist' and non-modal, in that only chunks of the real world are needed in its semantics. But this, too, evaporates, once one tries to give a situation-theoretic account of belief and being wrong. Significantly, [Barwise 1997; 'Information and Impossibilities'] eventually introduces some sort of worlds and modal languages into situation theory, though the resulting modal logics are not of the usual kind. Indeed, the more striking historical development has been the opposite, viz. the steady discovery of analogies between situation theory and modal logic. This has happend to such an extent that, in the mathematical framework of [Barwise & Moss], modal logic is the vehicle for developing a bisimulation-based situation theory! Indeed, the two paradigms seem compatible and congenial – and share a common cause.

In this section, we will emphasize two aspects of this (our treatment largely follows [van Benthem 2005]): (a) modal and first-order logics of dependence as an account of constraints, (b) merges between constraint-based models and dynamic epistemic logic. This section is more like a digression in this chapter, but it serves two real purposes. First, since both aspects of information are relevant in practice, we need to understand how epistemic logic and situation theory can live together in modeling actual phenomena. Next, juxtaposing these systems as accounts of information raises some interesting new questions, and allows us to see where both lie in the larger picture of logic today.

5.1 Correlation in state spaces

In a world of one-shot events, no significant information can flow. Constraints arise in situations with different 'states' that can be correlated. To make this more precise, consider two situations s_1, s_2, where s_1 can have some proposition letter p either true or false, and s_2 a proposition letter q. There are 4 possible configurations:

s_1: p, s_2: q	s_1: p, s_2: $\neg q$
s_1: $\neg p$, s_2: q	s_1: $\neg p$, s_2: $\neg q$

With all these present, one situation does not carry information about another, as p and q do not correlate in any way. A significant constraint on the total system arises only when we *leave out* some possible configurations. E.g., let the system be just:

s_1: p, s_2: q,	s_1: $\neg p$, s_2: $\neg q$

Now, the truth value of p in s_1 determines that of q in s_2, and vice versa. Stated in a formula with some obvious notation, we have the truth of the following constraint:

$$s_1: p \leftrightarrow s_2: q$$

But even a less constrained system with three instead of just two global configurations allows for significant information flow:

s_1: p, s_2: q,	s_1: $\neg p$, s_2: q,	s_1: $\neg p$, s_2: $\neg q$

Presence of p in s_1 still conveys the information that q in s_2, but absence of p does not convey information about s_2. Again in a formula, we have the implication:

$s_1 : p \rightarrow s_2 : q$

Contraposing the implication, absence of q in s_2 tells us about absence of p in s_1, but presence of q has no immediate informative value about s_1.

Thus, correlation between different situations amounts to restrictions on the total state space of possible simultaneous behaviors. The more 'gaps' in that state space, the more information there is in the system, to be used by potential observers. This view is similar to that of epistemic logic, where gaps in the set of worlds encode common knowledge. The bare bones of this setting are brought out by *constraint models*

$\boldsymbol{M} = (Sit, State, \boldsymbol{C}, Pred)$

with a set *Sit* of situations, a set *State* of possible valuations, a predicate *Pred* telling us which atomic predicates hold at which states, and crucially, a 'constraint relation' *C* telling us which assignments of states to situations are possible in the system.

5.2 *Modal logics of constraints and correlation*

These models suggest a logic. And as it happens, the simplest logic of constraint models is a modal one! Take a language with names x for situations (a tuple $\boldsymbol{x}$ names a tuple of situations), and atomic assertions $P\boldsymbol{x}$ for properties of or relations between situations. We take Boolean operations, plus a universal modality $U\phi$ ('ϕ is true everywhere'):

$P\boldsymbol{x} \mid \neg \mid \vee \mid U$

The semantic interpretation has obvious inductive clauses for the following notion:

$\boldsymbol{M}, \boldsymbol{s} \models \phi$ ϕ is true in global state $\boldsymbol{s}$ of model $\boldsymbol{M}$

In particular, we have that $P\boldsymbol{x}$ holds at $\boldsymbol{s}$ if the tuple of local states assigned by $\boldsymbol{s}$ to the tuple $\boldsymbol{x}$ satisfies the predicate denoted by P. This language defines basic constraints across situations such as $U(Px \rightarrow Qy)$. The resulting logic is classical propositional logic plus the modal logic *S5* for the universal modality U. But we can go one step further!

Intuitively, a situation x which satisfies p 'settles' the truth of p, plus all repercussions this has for other situations in the system. Define this new relation between global states:

$\boldsymbol{s} \sim_x \boldsymbol{t}$ *iff* $\boldsymbol{s}(x) = \boldsymbol{t}(x)$.

This generalizes to a relation $\sim_{\boldsymbol{x}}$ for sets or tuples of situations $\boldsymbol{x}$ by requiring equality of $\boldsymbol{s}$ and $\boldsymbol{t}$ for all coordinates in $\boldsymbol{x}$. Accordingly, there are modalities $[]_{\boldsymbol{x}}\phi$ for each such tuple, which say intuitively that the situations in $\boldsymbol{x}$ settle the truth of ϕ in the current system:

$\boldsymbol{M}, \boldsymbol{s} \models []_{\boldsymbol{x}}\phi$ iff $\boldsymbol{M}, \boldsymbol{t} \models \phi$ for each global state $\boldsymbol{t} \sim_{\boldsymbol{x}} \boldsymbol{s}$

This language expresses subtle properties of the information in the current model. One vivid metaphor here views situations as *agents*. Operators $[]_x$, $[]_{\boldsymbol{x}}$ then express what single situations or groups of them may be said to *know* on the basis of inspecting their own local properties. This epistemic interpretation of constraint models is strengthened by one more analogy. The tuple modalities $[]_{\boldsymbol{x}}$ involve an *intersection* of accessibility relations $[]_x$ for single situations x. This is like 'distributed knowledge' for groups in epistemic logic, describing what whole sets of agents may be said to 'know implicitly'.

Constraint models satisfy the following persistence properties for atomic facts:

$$\boldsymbol{Px} \rightarrow []_{\boldsymbol{x}}\boldsymbol{Px}, \qquad \neg \boldsymbol{Px} \rightarrow []_{\boldsymbol{x}}\neg \boldsymbol{Px}$$

This does not hold for all formulas. E.g., $[]_x Px \rightarrow []_y\, []_x Px$ is not valid, since accessible global states for x may change after a shift in the y coordinate. The extended modal constraint language has a decidable complete logic consisting of modal *S5* for each tuple modality, plus all axioms of the forms $U\phi \rightarrow []_{\boldsymbol{x}}\phi$, and $[]_{\boldsymbol{x}}\phi \rightarrow []_{\boldsymbol{y}}\phi$ whenever $\boldsymbol{y} \sqsubseteq \boldsymbol{x}$.

5.3 The larger picture: first-order logic of dependence

A natural alternative to our story of situations and constraints is in terms of *dependence*. Indeed, situations are like *variables x, y, ...* which can store values. A global state $\boldsymbol{s}$ is then a *variable assignment* in the usual sense: i.e., a function assigning an object to each variable. Now standard first-order logic has no genuine dependencies between variables. In any assignment $\boldsymbol{s}$, we can shift the value of x to some object d to obtain a new assignment $\boldsymbol{s}[x:=d]$, where all other variables have retained their $\boldsymbol{s}$-value. This is the reason why first-order logic typically has validities like commutation of quantifiers:

$$\exists x\, \exists y\, \phi \leftrightarrow \exists y\, \exists x\, \phi$$

The order of changing values is completely independent. But in many natural forms of reasoning, e.g., in probability theory, variables *x, y can* be dependent, in the sense that changes of value for one must co-occur with changes of value for the other.

Dependence of this sort can be modeled in a first-order setting ([van Benthem 1996, Ch. 9, 10]). Let a *general assignment model* be a pair $(\boldsymbol{M}, \boldsymbol{V})$ with $\boldsymbol{M}$ a first-order model with domain D and interpretation function I, and $\boldsymbol{V}$ a non-empty set of assignments on $\boldsymbol{M}$, i.e., a subset of the total function space D^{VAR}. The first-order language is interpreted as usual, now at triples $\boldsymbol{M}, \boldsymbol{V}, s$ with $s \in V$ – with the following clause for quantifiers:

$$\boldsymbol{M}, \boldsymbol{V}, s \models \exists x\, \phi \quad \text{iff} \quad \textit{for some } t \in \boldsymbol{V}\textit{: } s =_x t \textit{ and } \boldsymbol{M}, \boldsymbol{V}, t \models \phi$$

Here $=_x$ relates assignments identical up to x-values.

The analogy with our constraint language will be clear. Moreover, we also get new operators, such as *polyadic quantifiers* $\exists\boldsymbol{x}$ binding tuples of variables $\boldsymbol{x}$:

$$\boldsymbol{M}, \boldsymbol{V}, s \models \exists\boldsymbol{x}\ \phi \quad \text{iff} \quad \textit{for some } t \in \boldsymbol{V}\text{: } s =_{\boldsymbol{x}} t \textit{ and } \boldsymbol{M}, \boldsymbol{V}, t \models \phi$$

Here, $=_{\boldsymbol{x}}$ is identity between assignments up to values for all variables in $\boldsymbol{x}$. In first-order logic, $\exists xy\bullet\ \phi$ is just short-hand for $\exists x\exists y\phi$ or $\exists y\exists x\phi$ in any order. But in our new models, these expressions are no longer equivalent, as not all 'intermediate assignments' for x- or y-shifts need be present – and both are non-equivalent to $\exists xy\bullet$ as defined just now. One can also interpret single or polyadic *substitution operators* in their own right in general assignment models: $\boldsymbol{M}, \boldsymbol{V}, s \models [\boldsymbol{y}/\boldsymbol{x}]\ \phi$ iff $s[\boldsymbol{x}{:=}s(\boldsymbol{y})] \in \boldsymbol{V}$ & $\boldsymbol{M}, \boldsymbol{V}, s[\boldsymbol{x}{:=}s(\boldsymbol{y})] \models \phi$.

The logic of general assignment models is a decidable subsystem of standard predicate logic, which contains the valid first-order laws which hold even when variables may be correlated ([Németi 1996, van Benthem 1996, 2005]). Van Benthem 2005 shows how the above modal constraint logic can be faithfully embedded into the first-order logic of dependent variables, and also vice versa. Thus, we have the same topic in two different guises! Dependency is a major issue in foundations of first-order logic these days ([van Lambalgen 1996, van Benthem 1996, Abramsky 2005, Väänänen 2005]). Our analysis in the preceding passage shows that this fundamental theme is at the same time a move toward a general logic of information and constraints in the situation-theoretic sense.

Remark Beyond this decidable core logic, further axioms express special features of constraint models. E.g., the commutativity law $\exists x\ \exists y\ \phi \rightarrow \exists y\ \exists x\ \phi$ says that a *Diamond Property* should hold: If $s \sim_x t \sim_y u$, then there is another available assignment v with $s \sim_y v \sim_x u$. Imposing such special conditions on models makes them much like full function spaces, and the complete first-order logic (of *independence*) becomes *undecidable*.

5.4 Connections with epistemic logic and information as range

The basics of static and dynamic epistemic logic were explained in Sections 2 and 3.

Now, information as correlation and information as range seem different notions. The blinking dot on my radar screen has the information about an airplane approaching. But it does so whether or not I observe it. Whether I *know* that there is a blinking dot is an additional issue. It depends on whether I have observed the screen. Once I have made that observation S, and assuming I also know the right constraint $S \Rightarrow A$, I will indeed also know that there is an airplane A. To bring the two phenomena together, we need a *combined modal logic* of constraints and knowledge, first static, and then dynamic.

Consider a combined *epistemic constraint language* with syntax

$$P\boldsymbol{x} \mid \neg \mid \vee \mid U \mid [\,]_{\boldsymbol{x}} \mid K_i$$

Epistemic constraint models are of the form

$$\boldsymbol{M} = (Sit, State, \boldsymbol{C}, Pred, \sim_i)$$

with equivalence relations $\sim_i$ for each agent i. The right epistemic relations for our agents between global states depend on how we specify the scenario: they describe agents' access to the situational structure. The general logic of epistemic constraint models is a mere fusion of that for constraint models and a poly-*S5* epistemic logic of knowledge. We will mostly look at one agent, making the universal U behave like an epistemic K.

We now have a setting for disentangling correlation talk and range talk. Suppose that our model $\boldsymbol{M}$ satisfies the constraint $s_1{:}\ p \rightarrow s_2{:}\ q$. Then the agent knows this, as the implication is true in all worlds in $\boldsymbol{M}$. Now suppose the agent knows that $s_1{:}\ p$. In that case, the agent also knows that $s_2{:}\ q$, by the Distribution law of epistemic logic:

$$(K\ s_1{:}\ p \wedge K\ (s_1{:}\ p \rightarrow s_2{:}\ q)) \rightarrow\ K\ s_2{:}\ q$$

The converse requires more thought. The point is not that the agent already knows that $s_1{:}\ p$, but that, if she were to *learn* this fact, she would also know that $s_2{:}\ q$. In dynamic-epistemic terms, this reads $[!s_1{:}\ p]\ K\ s_2{:}\ q$. This formula is equivalent to the truth of the constraint – by the reduction axioms for dynamic-epistemic logic. Next, what do agents know about the informational content of specific situations x? If $[]_x\phi$ holds at some world $\boldsymbol{s}$, will the agent necessarily know this fact: $[]_x\phi \rightarrow \boldsymbol{K}\, []_x\phi$? Not so: $[]_x\phi$ can be true at some worlds, and false at others. What a situation x 'knows' need not be known to an external agent, unless this agent makes an observation about x. Thus, a combined modal-epistemic logic studies the interaction between our two senses of information.

Digression Here is one more analogy between constraint models and epistemic ones. In the literature (cf. [Fagin, Halpern, Moses & Vardi 1995]), epistemic worlds are often vectors of 'local states' for individual agents, and epistemic accessibility between worlds $\boldsymbol{s}, \boldsymbol{t}$ for an agent i is just component-wise equality $(\boldsymbol{s})_i = (\boldsymbol{t})_i$. This extends to $\sim_{\boldsymbol{i}}$ for groups of agents $\boldsymbol{i}$, just as we did in constraint models with the move from $\sim_x$ to $\sim_{\boldsymbol{x}}$. Thus, we can also use one uniform vector format for our combined models! Consider once again

$s_1{:}\ p,\ s_2{:}\ q$	$s_1{:}\ \neg p,\ s_2{:}\ \neg q$	$s_1{:}\ \neg p,\ s_2{:}\ \neg q$

Let some agent i have an accessibility structure indicated by the black dotted line:

$s_1{:}\ p,\ s_2{:}\ q$ ■■■■■■■■	$s_1{:}\ \neg p,\ s_2{:}\ \neg q$	$s_1{:}\ \neg p,\ s_2{:}\ \neg q$

We can bring this into vector format by adding the agent as a further component. It can be in one of two states, matching the 2 equivalence classes in this picture:

s_1: *p*, s_2: *q*, *i*: *state-1*	s_1: ¬*p*, s_2: ¬*q*, *i*: *state-1*
s_1: ¬*p*, s_2: ¬*q*, *i*:*state-2*	

The component-wise accessibility pattern is the same as in the preceding picture.

To summarize, we have seen that information as correlation and information as range co-exist happily inside one formal modeling, and that even in more than one way.

5.5 Events, information change, and correlation

Correlations and constraint models also have a dynamic aspect. We think of an evolving system in different global states that can change over time. The ground observation post remains in harmony with events on the mountain top over time. This temporal dynamics is there behind most of the key examples in situation theory, and it seems of interest to bring out this temporal dynamics explicitly in some formal language. And again, the simplest way of doing justice to these intuitions is by means of a modal formalism!

In what follows here, we will use a dynamic logic for this purpose, though a temporal one is quite feasible, too (cf. Parikh & Ramanujam 2002 on events and message passing). This time, we consider *dynamic constraint models*

$$\boldsymbol{M} = (Sit, State, \boldsymbol{C}, Pred, Event)$$

where events *e* are binary transition relations between global states. E.g., we may have had absence of a fire and a smoke signal, and then a combustion event takes place, changing the global state in our Mountain Top setting to *(smoke, fire)*. Our language combines the earlier constraint language with event modalities from dynamic logic:

$$P\boldsymbol{x} \mid \neg \mid \vee \mid U \mid [\,]_x \mid [e]$$

Here, we interpret the dynamic modality in the usual modal style:

$$\boldsymbol{M}, \boldsymbol{s} \models [e]\phi \quad \text{iff} \quad \text{for all } \boldsymbol{t} \text{ with } \boldsymbol{s}\, R_e\, \boldsymbol{t}: \boldsymbol{M}, t \models \phi$$

This language still describes constraints on the current state, but also what happens as the system makes some moves in the space of all its possible developments. This setting is close to DEL, as discussed in Section 4. Suppose for illustration that some single-situation system can have *2* states, and we do not know if it is in state *P* or ¬*P*. Now we perform an observation, and learn that it is in state *P*. This is not really an internal 'system event' in the preceding sense, as it affects some outside agent's view of where that system finds itself. We can treat such public observations *!P* in standard dynamic-epistemic style as changing the current constraint model ***M*** to its submodel ***M***|*P* leaving only those states that satisfy *P*. This distinction between system-internal events and external observation-events can be implemented straight-forwardly in the combined epistemic constraint models. The agent is trying to find out what the current state ***s*** is, though she also allows for the fact that ***s*** may change to ***t*** after system-internal events *e*.

The following table summarizes our basic perspectives so far:

	correlation	*range*
static	constraint logic	epistemic logic
dynamic	dynamic constraint logic	dynamic epistemic logic

As we have seen by now, these perspectives are not exclusive: they can be merged! Van Benthem 2005 gives a combined logic of interactions between all four components: correlations, and ranges containing information, and dynamic events modifying these – and even provides a joint embedding into the Guarded Fragment of first-order logic.

5.6 Co-existence

Our conclusion is that the two notions of information as range and as correlation are intimately related. Moreover, the two agendas naturally merge into one logical semantic view of information structure and dynamics. We can describe both the dependency-oriented information flow in a multi-component system, and the communication-driven oriented information flow through multi-agent systems accessing the former.

6 Information as Code: Syntax, Proof and Computation

Our discussion has been largely semantical in orientation so far, and it seems reasonably successful in signing up logical notions of information into one perspective. But within the logical tradition, there is another broad take on information structure and information processing, and it was already right there in our Introduction, viz. inference on formulas. The relevant part of logic is then *proof theory*, rather than model theory. Since this is such a vast area, we only discuss a few basic issues here, merely fleshing out a few relevant features of what might be called *information as code*. We will attempt nothing like a substantial exposition of proof theory (cf. [Troelstra & Schwichtenberg], [Feferman]). Indeed, the precise notion of 'information' behind the many calculi of proof in modern logic – and an optimal level of abstraction bringing it out explicitly –, seems even less established than that underlying logical semantics. An answer might lie in some form of abstract type theory making room for both classical logic, intuitionistic logic, and substructural logics ([Prawitz], [Belnap], [Restall]), but we will not go into this here. We refer to [chapter Abramsky] for some state-of-the-art thinking about these matters.

In Section 7 we will discuss, however, how proof-theoretic thinking links up with the above semantic views of information. That there must be some significant connection is clear from the *completeness theorems* which started off modern logic ([Gödel 1929]) by pointing out how, in major systems like first-order logic, a formula ϕ is true in all semantic models if and only if ϕ has a syntactic proof. Nevertheless, we have to confess to one more perplexity: the precise relationship between proof-theoretic and semantic views of information structure and information dynamics remains mysterious in many ways. There are some promising proposals, but nothing like a consensus in the field.

How can inference be informative? Let us start off with a somewhat surprising question, to which no definitive answer appears to be known in logic. Simply put, our earlier semantic theories of information have the following problem:

> *Semantic approaches do not account for the informative nature of* inference*!*

Natural tasks which we perform show a clear interplay of update through observation (or recollection, or whatever source) with more combinatorial inferential steps. When solving a puzzle, we do not just update information spaces, we also make appropriate *deductions* which highlight some important aspect of the solution, or at least, the road toward it.

You are throwing a party for three famous academics – a notorious social mine-field:

(a) John comes if Mary or Ann comes.
(b) Ann comes if Mary does not come.
(c) If Ann comes, John does not.

Can you invite professors in a way that respects this? In principle, you could update as in Section 2 here, starting with 8 possible sets of invitees, and then ruling out 3 by (a), 2 more by (b), and finally, using (c) to cut down to the only solution remaining. But in reality, you would probably do something like this (logicians' annotation included):

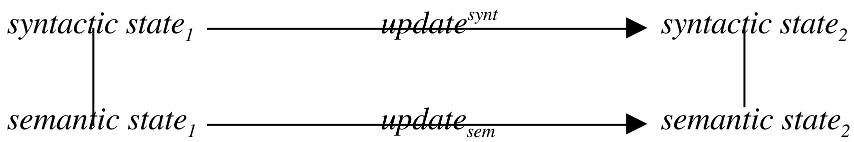
> By (c), if Ann comes, John does not. But by (a), if Ann comes, John does: a contradiction, so *Ann does not come*. But then, by (b), *Mary comes*. By (a) once more, *John comes*. Indeed, {John, Mary} satisfies all requirements.

There is a clear sense in which these successive inferential steps add information. Indeed, cognitive science tells us ([Knauff]) that much of what the brain does in cognitive tasks is a mixture of model inspection and inferential steps. But there is also a clear difficulty in saying in just which way the latter process *is* indeed informative. An inference does not shrink the current epistemic information range, and it does not add to what information the relevant constraint plus the current situation already 'contained'. It rather seems to increase information in a more combinatorial sense. Other versions of this problem are known in the philosophy of science. E.g., much has been written about the way deductive consequences of a physical theory, such as Einstein's deduction for the perihelion perturbation of Mercury from the General Theory of Relativity ([Fitelson]), can uncover startling new information which a scientific community did not have before.

Information in computation This problem is not unique to logic: it appears in several chapters of this Handbook. In particular, inference is much like computation, and one puzzles running through the chapter [Abramsky] is just how computation can provide information, and how abstract semantics can solve this. In a more standard mathematical framework, chapter [Adriaans] addresses this same issue of possible information growth through computation, but now in terms of numerical information measures such as Kolmogorov complexity. From the latter point of view, too, inference or computation can increase information, but only by a little, since a shortest description of the theorem prover generator plus the allotted running length suffices. These analogies show how the divide between more semantical and more code-based approaches to information occurs inside logic, too. In what follows we discuss a few strands in the latter direction.

Logical syntax In response to these observations, various answers have been proposed. For instance, Hintikka 1973 [Logic, Language Games, and Information] proposed a way inside first-order predicate logic for distinguishing levels of information in the syntactic analysis of a formula. First-order formulas ϕ describe an interplay of existence of objects with certain properties and excluding others, and these enumerations of object types can be unpacked level by level until the full quantifier nesting depth of ϕ. Inferences that can be made close to the original syntax contain only 'surface information', those requiring the full processing into normal from carry 'depth information'. Similar distinctions have been used for defining the classical philosophical 'analytic/synthetic' distinction as a matter of first-order logic. All this is clever, and it suggests that logical syntax might have a role in defining its own natural notion of information content: one where syntactic manipulation can increase information. But so far, no coherent and generally useful notion of syntax-based information has evolved – partly because we have no general theory of syntax, abstracting away from details, which would provide the vehicle for this.

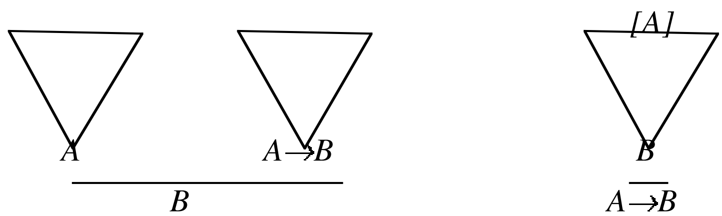
Two-level syntax-semantics formats One type of proposal, made in various guises in recent years, is much simpler – and it often comes in the form of an amalgam between semantic and syntactic views. Suppose that the information state of agents really has two parts: (a) the current range of models or possible worlds, and (b) *some current set of sentences* true in these models which the agent has available right now. Like the worlds in epistemic logic, and the situations of Situation theory, the syntax domain, too, may be structured – say by an ordering of formulas as to relative importance ([Ryan], [Gärdenfors]). Then we can think of informational processes as operating on two levels:



Our earlier range-changing updates may then modify the lower level, while inference or other syntactic procedures modify the upper syntax component of the current state, and hence the information ' immediately present' to the agent. This process can even be described in the same sort of dynamic logic format used for epistemic logic in Section 3. Examples of this sort of approach are Dung [on dynamic logic of inference steps], Wassermann [on belief revision over syntactic knowledge bases], Ryan [on structured default theories], Gabbay, Kempson & Meyer Viol [on dynamic logic of syntax], and Jago [on fine-grained process models of Horn clause inference, treating the universe of Herbrand models (sets of literals) as an inference machine whose moves lead to richer information stages. Also in this tradition is [Parikh 2007; "Sentences, Propositions and Logical Omniscience, or What does Deduction tell us?", CUNY 2007] proposing an abstraction out of this which should again explain how deduction can be informative.

Finally, two- and many-level views are coming up at the interface of logic and *cognitive science* ([Castelfranchi & Lorini 2007]) to do justice to the complex array of cognitive processes taking place simultaneously when humans actually process information. Philosophers and mathematicians often invoke intuitions or metaphors in this vein to motivate their theories, but there are also systematic and intriguing facts of the matter which may become increasingly relevant to logical theorizing ([van Benthem 2007]).

Proofs and information dynamics But the most sophisticated paradigm bringing out information flow in inference and computation is logical *Proof Theory* ([Schwichtenberg & Troelstra], [Girard]). Proofs may be seen as ordered structures of sentences with their inferential links indicated, and hence they are a natural candidate for representing information structure at a syntactic level. For instance, a proof system like *natural deduction* is an elegant calculus for putting together , linking up and re-packaging pieces of evidence – witness proof rules like those for the conditional:



The view of mathematical proof calculi as general systems for *combining evidence for claims* has been put forward forcefully by authors such as [Martin–Löf], [Prawitz]. Moreover, there are natural notions of proof equivalence, e.g., through the valid identities of the lambda calculus and related abstract category-theoretic notions of equivalence ([Lambek & Scott]). The resulting normal forms provide an attractive abstract level of information structure beyond brute details of syntax, and in the setting of linear logic, [Girard] has emphasized how normalization steps may be seen as communication moves, and hence as a form of information flow. We will not elaborate the most wide-spread and elegant natural deduction, type-theory, and category-based proof-theoretic paradigm in this chapter, but we refer to [Abramsky] for an approach very much in that spirit.

For our present purposes, suffice it to say that one can think of inference steps as updating proof structures, and hence as information transformers at some finer level of current information structure. This seems in line with the *intuitionist* slant of many proof theories ([Dummett], [Troelstra & van Dalen], [Sundholm]), where one thinks intuitively of proof as being performed by agents investigating objects and moving through various information stages in the course of an enquiry. This interpretation has even been given an attractive dynamic interpretation in the *dialogue games* of [Lorenzen], where proof rules correspond to game moves, and proofs themselves are strategies for players involved.

Section 7 Semantics and Proof: merging internal and external evidence

The proof-theoretic story of information, though suggestive, raises some obvious objections. Mathematical proof, however elegant, hardly seems the paradigm for all the kinds of evidence that humans manipulate, or for the ways in which they do so. Clearly, we are not just playing what [Hintikka 1973] called 'indoor games' of proof or argumentation. Some impeccable evidence rather comes from our senses, some from being told, and so on. And these phenomena were described rather naturally in our earlier semantic accounts of information. Indeed, it seems highly implausible that all significant information flow can be understood as proof or computation steps. Intuitively, the latter rather seem to describe the *internal elucidation* of information already obtained, while the former handle the *external uptake* of incoming new information.

But stated in this way, a form of harmony suggests itself! One way of having the best of both worlds here would be to combine proof-theoretic viewpoints with the earlier semantic viewpoints of this chapter into one broader calculus of knowledge and evidence:

Range and evidence: a case of ∀ versus ∃ Indeed, proof-theoretic and semantic perspectives on information are not in conflict: they complement each other. But how?

Let us first high-light a logical difference. As we saw in Section 2, epistemic logic of information as range revolves around a *universal quantifier*: $K_i\phi$ says that

ϕ is true in all situations *agent i considers as candidates for the current s.*

But there is also an existential quantifier in the famous traditional account of knowledge that ϕ as 'justified true belief. The first phrase in this definition really says that

there exists a justification for the formula ϕ.

In this sense, knowledge consists in having evidence for a proposition: sometimes, on an exceptionally clear Cartesian day, even a mathematical proof. But as we saw, the two views co-exist in logic! The completeness theorem for first-order logic establishes an equivalence between validity (an $\forall$-type notion) and provability (an $\exists$-type notion).

Combined calculi To have *both views together* then, [van Benthem 1993] proposed merging standard epistemic logic with a 'calculus of evidence'. Candidates for this range from *intuitionistic logic* with binary type-theoretic assertions of the form

x is a proof for ϕ.

([Dummett], [Troelstra & van Dalen], [Tennant]) to the much more general 'labeled deductive systems' of [Gabbay 1996], which were designed as a general calculus of statements *x: ϕ,* where the label *x* can stand for a proof, an observation, or indeed any kind of evidence from any source, possibly even with further information attached about reliability, probability, and so on, of the support that *x* offers to ϕ.

Even closer to the epistemic logics of Sections 2,3, the 'Logic of Proofs' of [Artemov 1994, 2005] is a modal language with evidence-labeled assertions $[x]\phi$. Then, e.g., the ubiquitous (though philosophically controversial; [Egre 2005]) epistemic Distribution Axiom $K_i(\phi \to \psi) \to (K_i\phi \to K_i\psi)$ becomes the more informative statement that

$$[x](\phi \to \psi) \ \& \ [y]\phi \ \to \ [x\#y]\psi,$$

where # is some natural *sum operation* on proofs, or pieces of evidence generally. Incidentally, van Benthem 2005 shows how this also works for contextualist views of information – as discussed in chapter [Dretske]. In that case the crucial law is

$$[c_1]\, K_i(\phi \to \psi) \ \& \ [c_2]\, K_i\phi \ \to \ [c_1\# c_2]\, K_i\psi,$$

where # is now an operation of 'context merge' reminiscent of situation theory.

Merging internal and external dynamics In line with our proposed division of labour between *external uptake* and *internal elucidation* of information, a more ambitious goal would be a logical calculus which combines the dynamics of *internal* proof steps with that of *external* observation steps such as the public announcements of Section 3. This would combine the 'dynamics of observation' with the 'dynamics of elucidation'. Both seem essential facts of human cognitive functioning, witness our Introduction. One might think this is exactly what is being provided by the dynamic epistemic logics of Section 3, or the situation-theoretic calculi of Section 4. But neither of these provide an explicit description of the dynamics of information corresponding to individual inference steps. Some first attempts in the latter mode are the dynamic logics of [Dung] [Jago 2006], which operate on more finely-grained universes of 'proof stages'.

A deeper identity? All this shows that evidence based $\exists$-type accounts and range-based $\forall$-type accounts of information can live together. It does not answer the deeper question whether there is some deeper identity between evidence combination in an epistemic sense and that of modern proof theory. Stated differently, *proof theory* and *dynamic logics* often address very similar issues and provide solutions, whether in process theory or information dynamics, but the precise analogy remain a bit of a mystery. For instance, [Van Benthem 1997] unifies both in terms of a ternary notion of combination, with rules

(a) $x : A, \quad y : A \to B, \; Rz, xy \;\; \vdash \;\; z : B$

where *Rz, xy* is some *ternary* condition relating *z, x, y*.

The atom *Rz, xy* can be read as '*z* is the sum of the information pieces *x, y*', '*z* results from composing the transitions *x, y*', '*z* results from applying the function *x* to argument *y*', etc. Principle (a) is the abstract analogue of the earlier Modus Ponens in natural logic. Here is the matching principle for the dual rule of Conditionalization:

(b) $\tau : X, \quad x : A, \; Rz, xy \;\; \vdash z : B$ implies $\tau : X \; \vdash \; y : A \to B$

But so far, this has remained an abstract format without a convincing intuitive unification.

Information that and information how We conclude with one final illustration how merging semantic and proof-theoretic perspectives on information in logic can be of interest. Proof theory has a clear *algorithmic* aspect, which we have downplayed in this chapter (though this whole Handbook has quite a lot of it). At least so-called constructive proofs are at the same time ways of 'seeing' and methods for realizing the *conclusion*. This dual aspect to inference: proving propositions and delivering methods, has been around ever since the so-called 'constructions' in Euclid's *Elements*. It is even more pronounced in type-theoretic approaches where proofs are at the same time definable function terms in some mathematical universe through the Curry-Howard isomorphism. This same insight underlies the treatment of information structure in chapter [Abramsky]. Moreover, this duality ties in with a more general famous distinction in philosophical epistemology between *knowledge that* and *knowledge how* ([Ryle], [Gochet]).The latter is about procedures establishing truth, and more general skills.

In our view, a similar propositional – procedural distinction holds for the notion of information. Information is not just about *what things are like* in some situation, but also *how to get thing things done*. Merging proof-theoretic and semantic systems of logic seems an excellent way of doing justice to that intuitive feature.

Conclusion Semantic and proof-theoretic views of logical information are different and they operate on different sorts of information structure. Nevertheless, the development of logic shows that they live in close harmony – and indeed, a merge of both perspectives may be needed to really understand the notion of information in its entirety. We have not really proposed one best way, let alone a best calculus, for doing this, and leave that as a future challenge to our readers. In leaving matters at this stage, we console ourselves with the thought that this may just reflect the general complementarity of more semantic and more algorithmic views of information which runs throughout this entire Handbook.

8 Further Topics and Broader Outreach

Our presentation has focused on a few major senses of information in logic, and systems showing how they work. This is by no means an exhaustive survey of all issues, and one or more alternative chapters could have been written mapping out the territory differently. Of the omissions that weigh most on our conscience, we mention the following:

Information content and methodology Carnap's work in the 1950s. Connections to issues of evidence and confirmation in the philosophy of science.

Intuitionism and constructive mathematics Early logics of investigations through information stages and constructive processes: Beth models, Kripke models. Topological models of information: [Tarski], [Scott], [Vickers].

Modal and substructural logic of information stages Modal logics: [van Benthem 1994], [Sequoiah-Grayson 2007]. Substructural logics: [Dunn], [Belnap], [Restall].

Belief revision and syntactic belief bases [Wassermann], [Rott], [Tamminga].

Other information carriers Symbolic versus graphic, Diagrammatic reasoning [Barwise & Etchemendy], [Kerdiles]. Spatial patterns: [Handbook of Spatial Logics].

Information and context Situation theory and the McCarthy – Trento School.

Logic and probability Inductive logic, Probability theory, Merged probabilistic epistemic logics. Cf. [chapters Grunwald & Vitanyi, Topsoe & Harremoes, Adriaans].

In formation and computation Logic and computer science. Cf. [chapter Abramsky].

Information and dialogue Lorenzen dialogues, multi-agent update and belief revision.

Information and general interaction Game theory, economics [chapter Walliser], social sciences [chapter Devlin & Rozenberg].

How humans really pick up information Cognitive science [chapter Boden]

The last word for: philosophy Anything we forgot to say about epistemology philosophy of language, philosophy of science, etc.

9 Conclusion: Logic as Information Theory

This chapter has painted a picture of logic as a theory of information. This is not at all the standard view, let alone the standard self-image, of the discipline, but it does seem a legitimate perspective. We have woven a story which puts many logical notions and results into one new line, but one rather different from a standard textbook presentation. Here are our major conclusions, or rather: our claims, in doing so. First, to get the real picture of logical information, one needs to address the *statics and dynamics in parallel*, with intertwined accounts of information structure and the dynamic processes which manipulate the latter. Next, there are many such processes, which need not all reduce to one primitive, and hence logical views of information come in complementary kinds. First, we re-interpreted epistemic logic as an information theory of *range, knowledge*, and *observation-based update*. Like classical information theories, this is not just a model of information, but also a calculus for computing information flow, witness the concrete procedures for model update, and the axiomatic dynamic-epistemic laws governing this. We also emphasized the essential role of agents, and the ways in which *agents take information*: alone, but much more importantly, in *interaction* with others. Information is studied best in a setting of many agents, communication, and perhaps even interactive games. Next, we high-lighted another major perspective on information, viz. the situation-theoretic account of *correlations and dependence* between different parts of distributed systems. This provides a much richer view of how agents can actually access information, and why it is there for them to pick up and communicate in the first place. Our further claim is that this situation-theoretic perspective is not in conflict with the epistemic one, but rather forms its natural complement. But next, logic offers another major view of *information as code* in the form of proof systems, and other syntax-based calculi. Our view is that this is another informational process, and we propose a distinction (which is very common really) between the *uptake of external information* and the *internal elucidation* of information received. Our claim is that the two can live in harmony, – and that indeed, they can be merged in many interesting ways. Finally, at this stage, we reach the border line with information processing as computation, and more quantitative views of information transmitted by this and other dynamic processes.

At this stage, we hand over to the other chapters of this Handbook, leaving the question whether a Grand Unification is desirable, or even possible, to the editors.

10 References (still to be added)

J. Barwise & J. van Benthem, P. Lindström, information across situations.

E. Lorini & C. Castelfranchi, 'The Cognitive Structure of Surprise: looking for basic principles, *Topoi* 2007.

F. Giunchiglia & L. Serafini, context logics.