Modeling Background Conditions

Definition 19.6.

1. A *background condition* B is a function with domain some set P of input observables, taking real numbers as values. It is required that $B(i)$ be a number in the range of the ith observation function of the state space. The domain P is called the set of *parameters* of B.
2. A *state* σ satisfies B if $\sigma_i = B(i)$ for each $i \in P$. A *token* s *satisfies* B if $state(s)$ satisfies B.
3. The set of background conditions is partially ordered by inclusion: $B_1 \leq B_2$ if and only if the domain of B_1 is a subset of that of B_2 and the two functions agree on this domain.

Every nonempty set of background conditions has a greatest lower bound under the ordering $\leq$. Also, if a state or token satisfies B, then it clearly satisfies every $B_0 \leq B$. The empty function is the least background condition: It imposes no conditions on states.

In working with a given state space, one assumes one has a fixed background condition B, that one is only concerned with tokens that satisfy this background condition, and hence that all computations and inferences take place relative to that background condition. We can make this precise as follows:

Given a set Q of input observables, define $\sigma \equiv_Q \sigma'$ if $\sigma_i = \sigma'_i$ for all $i \notin Q$. This is an equivalence relation on states. We say that a type $\alpha \subseteq \Omega$ is *silent on* Q if for all states $\sigma, \sigma' \in \Omega$, if $\sigma \equiv_Q \sigma'$ and $\sigma \in \alpha$, then $\sigma' \in \alpha$. We say that α is silent on an input observable i if it is silent on $\{i\}$. Finally, a type α is *silent on* B if α is silent on the set P of parameters of B. If one is given a premise or purported conclusion that is *not* silent on an observable i, it is safe to assume i is *not* a parameter of the system. Put another other way, *if we are reasoning about an observable* i*, then* i *must be either an explicit input or output of the system.*

Example 19.7. In our heating system example S_{hs}, the types α_1, α_2, and β are silent on the parameters, as can be seen from their definitions. The type α_3, by contrast, says that $\sigma_3 = 0$ and so is not silent on σ_3. The natural background condition B in the winter is $\sigma_3 = \sigma_4 = \sigma_5 = 1$, representing the case where the power switch is on, the vents are not blocked, and the setting is on "heat." (This is our informal way of indicating the background condition B with domain $\{3, 4, 5\}$ and constant value 1.) In the summer the default background condition is $\sigma_3 = \sigma_4 = 1$ and $\sigma_5 = -1$.

Definition 19.8. Given a background condition B and a set Γ of types, the *weakening of* B *by* Γ, written $B\restriction\Gamma$ is the greatest lower bound (in the $\leq$ ordering) of all $B_0 \leq B$ such that every type $\alpha \in \Gamma$ is silent on B_0.

When $\Gamma = \{\alpha\}$, we write $B\restriction\alpha$ for $B\restriction\Gamma$. The function $B\restriction\alpha$ models the intuitive idea of dropping background assumptions of B on the parameters that are critical to the content of the type α. If the input observables of our state space are independent, then this definition is guaranteed to behave the way one would hope.

Theorem 19.9. *Let* Γ *be a set of types. Assume the input observables of the state space* S *are independent. Then each type in* Γ *is silent on* $B\restriction\Gamma$*, and this is the greatest such background condition* $\leq B$.

Proof. This follows easily from the following lemma, of independent use in computing $B\restriction\Gamma$. □

Lemma 19.10. *If the input observables of the state space* S *are independent, then a type* α *is silent on* Q *if and only if it is silent on each* $i \in Q$*. Hence, under these conditions, there is a largest set* Q *of input observables such that* α *is silent on* Q.

Proof. The left to right half of the first claim is trivial. The converse claim is proved by induction on the size of Q. The second claim follows from the first by taking Q to consist of all the input observables i such that α is silent on i. □

Given a background condition B with parameters P, the lemma tells us that $B\restriction\alpha$ is simply the restriction of B to the set of input observables $i \in P$ such that p is silent on i.

Example 19.11. The inputs of our example S_{hs} are independent so the lemma and theorem apply. The type α_3 is silent on each observable except for σ_3. Hence if B is $\{\sigma_3 = \sigma_4 = \sigma_5 = 1\}$, then Weakening B by α_3 gives $B\restriction\alpha_3 = \{\sigma_4 = \sigma_5 = 1\}$.

Relativizing to a Background Condition

Each background condition B determines a subspace $S_B \subseteq S$ as follows.

Definition 19.12. Let B be a background condition for the state space S. The *relativization* S_B *of* S to B is the subspace of S whose states are those states σ

of S that satisfy the background condition B and whose tokens are the tokens that satisfy B.

By relativizing, the output equations are simplified because the parameters become constants in the equations in S_B.

Example 19.13. In our example, the output equations simplify to

$$\sigma_6 = pos(sg(\sigma_1 - \sigma_2))$$
$$\sigma_7 = \begin{cases} 80 & \text{if } \sigma_6 = +1, \\ \sigma_2 & \text{otherwise.} \end{cases}$$

Recall the correspondence between subspaces of S and S-logics from Lecture 16. Because $S_B \subseteq S$, this gives us a $\mathrm{Log}(S_B)$ with $\mathrm{Log}(S) \sqsubseteq \mathrm{Log}(S_B)$. We call this the *local logic* $\mathrm{Log}(S_B)$ *supported by the background condition* B.

Proposition 19.14. *For each background condition* B*, the local logic* $\mathrm{Log}(S_B)$ *supported by* B *is the logic on* $\mathrm{Evt}(S)$ *given by the following:*

1. *The* $\mathrm{Log}(S_B)$*-consistent states* σ *are those satisfying* B*.*
2. *If* Γ, Δ *are sets of types of* S*, then* $\Gamma \vdash_B \Delta$ *if and only if for every state* σ *satisfying* B*, if* $\sigma \in p$ *for all* $p \in \Gamma$*, then* $\sigma \in q$ *for some* $q \in \Delta$*.*
3. *The normal tokens are those tokens satisfying* B*.*

Proof. The proof of (1) follows immediately from the definition and (2) and (3) follow from (1) and Theorem 16.6. □

Because $\mathrm{Log}(S) \sqsubseteq \mathrm{Log}(S_B)$, we know that $\mathrm{Log}(S_B)$ typically has more constraints but fewer normal tokens than $\mathrm{Log}(S)$, as one would expect. The logic $\mathrm{Log}(S_B)$ is not in general sound, because there may well be tokens not satisfying the background condition B. Indeed, $\mathrm{Log}(S_B)$ is sound if and only if every token satisfies B. $\mathrm{Log}(S_B)$ is complete if and only if every state that satisfies B is the state of some token.

Example 19.15. In our running example, the sequent $\alpha_1, \alpha_2 \vdash \beta$ is a constraint in the logic $\mathrm{Log}(S_{\sigma_3=\sigma_4=\sigma_5=1})$ with background condition $\sigma_3 = \sigma_4 = \sigma_5 = 1$. (To see this we need only show that if σ is a state with $65 \leq \sigma_1 \leq 70$ and $\sigma_2 = 58$, then $\sigma_6 = 1$ and $\sigma_7 > \sigma_2$. Using the state equations for $S_{\sigma_3=\sigma_4=\sigma_5=1}$ displayed above, we calculate that $\sigma_1 - \sigma_2 > 2$ so $sg(\sigma_1 - \sigma_2) = 1$. Hence $\sigma_6 = 1$ and $\sigma_7 = 80$, in which case $\sigma_7 > \sigma_2$, as desired.) On the other hand this constraint does not hold in the full logic $\mathrm{Log}(S_{hs})$, as can be seen by looking at, say, the state $\sigma = \langle 67, 58, 0, 1, 1, 0, 58\rangle$. This state satisfies the defining equations of our large space; it satisfies α_1 and α_2 but not β.

The following shows that as the background conditions increase in restrictiveness, so do the associated logics.

Corollary 19.16. *If* $B_1 \leq B_2$*, then* $\mathrm{Log}(S_{B_1}) \sqsubseteq \mathrm{Log}(S_{B_2})$*.*

Putting this the other way around, if the background conditions decrease in strength, so do the associated logics. As remarked earlier, in reasoning about a system we expect to be given information that is silent on the parameters of the system, because when we are given information that is not silent on some input observable, it is no longer a parameter but an explicit input of the system. In particular, if we are given explicit information about the value of some observable, these observables cannot be parameters, which means that the background is weakened according to the new information. But weakening the background weakens the logic. Thus additional information that is not silent on the parameters weakens the logic.

Example 19.17. The information in α_3 is about the observable σ_3, the power to the system. Not only is α_3 not silent on the background condition B of the first sequent, it directly conflicts with it. Thus the natural understanding of the claim that $\alpha_1, \alpha_2, \alpha_3 \vdash \beta$ is as being relative to the weakened $B\lceil\alpha_3$. Relative to the local logic $\mathrm{Log}(S_{B\lceil\alpha_3})$, this is not a valid constraint. Indeed, in this logic a routine calculation similar to the above shows that $\alpha_1, \alpha_2, \alpha_3 \vdash \neg\beta$, as desired.

We summarize the above discussion by putting forward a pragmatic model of the way people intuitively reason against background conditions. First we present the following definition.

Definition 19.18. Γ *strictly entails* Δ relative to the background condition B, written $\Gamma \Rightarrow_B \Delta$, if the following four conditions hold:

1. $\Gamma \vdash_{\mathrm{Log}(S_B)} \Delta 0$;
2. all types in $\Gamma \cup \Delta$ are silent on B;
3. $\bigcap \Gamma \neq \emptyset$;
4. $\bigcup \Delta \neq \Omega$;

The first two conditions have been extensively discussed above. The third and fourth are not important for this discussion, but we include them for the sake of completeness because they do capture intuitions about the way people reason.

The third condition is that the information in the sequent is consistent in that there should be some possible state of the system compatible with everything in Γ. The last condition is that the information in sequent is nonvacuous in that not every possible state of the system should satisfy Δ.

Our observations can now be put as follows:

1. The consequence relation $\Gamma \Rightarrow_B \Delta$ is a better model of human reasoning against the background condition B than is $\Gamma \vdash_B \Delta$.
2. The relation $\Gamma \Rightarrow_B \Delta$ is monotonic in Γ and Δ, but only as long as you weaken using types that are silent on B and do not make the sequent hold for trivial reasons.
3. If one is given a type α that is not silent on B, the natural thing to do is to weaken the background condition B by α, thereby obtaining $B\lceil\alpha$.
4. But $\Gamma \Rightarrow_B \Delta$ does not entail $\Gamma, \alpha \Rightarrow_{B\lceil\alpha} \Delta$ or $\Gamma \Rightarrow_{B\lceil\alpha} \Delta, \alpha$.

Logic Infomorphisms and Nonmonotonicity

A reasonable objection to our presentation so far would be to point out that we have systematically exploited an ambiguity by using "$\alpha_1, \alpha_2, \alpha_3$" and "$\beta$" for both symbolic expressions (like the English sentences where they were introduced at the start of this lecture) as well as for the "corresponding" state space types, which are, after all, sets of states, not exactly the sort of thing people reason with in ordinary life. We need to discuss this relationship between the symbolic expressions and said types.

The solution, of course, is our notion of a logic infomorphism. To see what this has to do with our problem, let us return to our example and set up a symbolic Boolean classification $\boldsymbol{A}$. The types of $\boldsymbol{A}$ are symbolic expressions $\varphi, \psi, \eta, \ldots$, either the English expressions used in the introduction and Boolean combinations of them or some sort of more formal counterparts of the sort used in an elementary logic course. We use $\varphi_1, \varphi_2, \varphi_3$, and ψ for the premises and conclusions of our classification. The tokens of $\boldsymbol{A}$ are instances of Judith's heating system. Let $\boldsymbol{B}$ be the classification Evt($\boldsymbol{S}$) associated with the state space $\boldsymbol{S}$ as above. The map $\varphi_i \mapsto \alpha_i$, and $\psi \mapsto \beta$, extended in the natural way to the Boolean combinations, defines a map $f^\wedge$ from the types of $\boldsymbol{A}$ to the types of Log($\boldsymbol{S}$), whereas the identity map on tokens can be thought of as a map of the tokens of Log($\boldsymbol{S}$). This gives us a classification infomorphism from $\boldsymbol{A}$ to the classification Evt($\boldsymbol{S}$),

$$f : \boldsymbol{A} \rightleftarrows \boldsymbol{B}.$$

We know that any infomorphism $f : \boldsymbol{A} \rightleftarrows \boldsymbol{B}$ and local logic $\mathfrak{L}$ on $\boldsymbol{B}$ gives rise to a $\sqsubseteq$-largest local logic $f^{-1}[\mathfrak{L}]$ on $\boldsymbol{A}$ such that f is a logic infomorphism from this logic to $\mathfrak{L}$. This gives us a host of logics on the classification $\boldsymbol{A}$. There is the logic $f^{-1}[\text{Log}(\boldsymbol{S})]$, but also, for each subspace of $\boldsymbol{S}$, there is the inverse image of the logic associated with this subspace. In particular, for each background condition B, there is a local logic $f^{-1}[\text{Log}(\boldsymbol{S}_B)]$. Let us write these logics as $\mathfrak{L}$ and $\mathfrak{L}_B$, respectively, and write $\Gamma \vdash \Delta$ and $\Gamma \vdash_B \Delta$ for their entailment relations.[2]

The logic $\mathfrak{L}$ is sound, that is, all its tokens are normal. The constraints of this logic are just those sequents of $\boldsymbol{A}$ whose validity is insured by our state-space model. The logics of the form $\mathfrak{L}_B$, however, are not sound; the normal tokens of $\mathfrak{L}_B$ are those that satisfy the background condition B. For example, the normal tokens of $\mathfrak{L}_{\sigma_3=\sigma_4=\sigma_5=1}$ are those in which the power is on, vents are unblocked, and the controls are in heating mode.

The Ineffability Principle

In the model given in the previous section, a background condition B could be packaged into a type γ_B, namely,

$$\gamma_B = \{\sigma \in \Omega \mid \sigma \text{ satisfies } B\}.$$

This type could, in principle, be made explicit as an additional premise of an inference. But this type will typically lie outside of the range of $f^\wedge$. Put differently, *there is in general no way to capture the background condition B by a symbolic type of* $\boldsymbol{A}$. We call this the "ineffability principle" because it seems to model the fact that it is seldom, if ever, possible to say exactly what background assumptions are in force when we reason in ordinary life.

Turtles All the Way Down

Having seen the relationship between state spaces, local logics, and nonmonotonicity, let us return to state spaces.

As a start, let us note that if $\boldsymbol{S}$ is a real-valued state space of dimension n and B is a background condition on k parameters, then $\boldsymbol{S}_B$ is isomorphic to a real-valued state space $\boldsymbol{S}_B^*$ of dimension $m = n - k$. The isomorphism is the identity on types whereas on tokens it projects $\langle\vec{\sigma}_i, \vec{\sigma}_p, \vec{\sigma}_o\rangle$ to $\langle\vec{\sigma}_i, \vec{\sigma}_o\rangle$. (In our example, the new space would have as states those 4-tuples $\langle\sigma_1, \sigma_2, \sigma_3, \sigma_4\rangle$ such that $\langle\sigma_1, \sigma_2, 1, 1, 1, \sigma_3, \sigma_4\rangle \in \Omega$.) The move from the

[2] We would actually propose combining this proposal with the ideas in Lecture 18 in a full development.

n-dimensional S to the m-dimensional S^*_B results from setting k parameters as dictated by B.

Now let us look at it the other way around. Suppose we had begun with an m-dimensional state space S and later learned of k additional parameters that we had not taken into account. Letting $n = m + k$, we would then see our space S as projecting onto a subspace S'' of an n-dimensional state space S'. Because $S'' \subseteq S'$, taking the associated local logics and recalling the order reversal that takes place, we have $\mathrm{Log}(S') \sqsubseteq \mathrm{Log}(S'')$. That is, the logic associated with the new state space is weaker, in that it has fewer constraints, but more reliable, in that it has more normal tokens, which is just what we have seen.

Example 19.19. Suppose instead of the information α_3, we had been faced with the following new information:

(α_4) The gas line has been broken and there is no gas getting to the furnace.

We would like to get $\alpha_1, \alpha_2, \alpha_4 \vdash \neg\beta$, but we do not. In fact, the type α_4 does not even make sense in our state space, because the gas pressure and gas line have not been taken into account in our seven-dimensional state space S_{hs}. To take these into account, we would need to see S as isomorphic to a subspace of a nine-dimensional state space S', one where the additional observables are the state of the gas line and the gas pressure. Then as we have seen, we would have $\mathrm{Log}(S') \sqsubseteq \mathrm{Log}(S)$.

One should not think that there is ever an "ultimate" completely perfect state space. In general, it seems it is almost always possible in real-life systems to further refine a state-space model of some system by introducing more observables. That, of course, is what we mean by saying it is turtles all the way down.

The example makes an additional point, though. Once we introduce the new observables, we have also implicitly changed the set of tokens under consideration. Our tokens were just instances of Judith's heating system, something located entirely in Judith's house. But the gas lines run from the house out under the streets of Bloomington. Our tokens have greatly expanded. Each of our old tokens s is part of a token $s' = f^{\vee}(s)$ in S'. In other words, our isomorphism is no longer the identity on tokens. When we look at this in terms of the associated classifications $\mathrm{Evt}(S')$ and $\mathrm{Evt}(S)$, what we have is an infomorphism $f : \mathrm{Evt}(S') \rightleftarrows \mathrm{Evt}(S)$ that is not the identity on tokens. Rather, it takes each token to a richer token. And again, there typically seems to be no end to this. As we reason in greater detail, there seems no end to the richness that may have to be considered.[3]

19.3 The Frame Problem

We view the frame problem as the problem of specifying the immediate consequences of a basic action a taken by an agent, what changes and what doesn't change. Given a state space S as above, following the example in Lecture 3, we define a new state space S_{Act}. Its tokens consist of basic actions. We assume that for each such act a there are two tokens $\mathrm{init}(a)$ and $\mathrm{final}(a)$ of S. We take as states the pairs $\langle\sigma_1, \sigma_2\rangle$ of states of S. We say that a token of S_{Act} has state $\langle\sigma_1, \sigma_2\rangle$ if $\mathrm{init}(a)$ has state σ_1 and $\mathrm{final}(a)$ has state σ_2. Now any sort of change to the circuit can be modeled in $\mathrm{Evt}(S)$.

The frame problem and the problem of nonmonotonicity interact with one another, so one would hope that the machinery developed here would help with the frame problem. Suppose that the temperature is seventy-two degrees and the thermostat is set at sixty-eight degrees, so no hot air is coming out of the vents. But now suppose we are told that the temperature drops to sixty-eight degrees. It seems we want to conclude that the furnace comes on and hot air comes out of the vents. But of course this would not be valid if, in the meantime, the setting was changed to the summer setting. Our proposal for handling the frame problem in this context is quite simple: described actions only involve changes to output observables and to input observables about which the descriptions are not silent. Otherwise one has to weaken the background as before.

19.4 Conclusions

In spite of their ubiquity in science and applied mathematics, state spaces as models for human reasoning have been largely ignored. But there are a couple of reasons why such a move might prove fruitful.

Feasibility

In those applications where the equations

$$\sigma_o = F_o(\vec{\sigma_i})$$

[3] This is not a new observation, it has been known to those working in nonmonotonic logic for a decade or so. The point here is just to see how this realization fits into the picture presented here.

really can be computed, the state-space/local-logic approach might give us an interesting alternative to traditional theorem-proving methods, one where fixed numerical calculations could be used in place of more symbolic approaches. This is the antithesis of Pat Hayes' thesis in his famous manifesto (Hayes, 1985). The proposal here is to *exploit* the sciences in modeling commonsense reasoning rather than replace it with a parallel symbolic framework.

In this regard, Tarski's decision procedure for the real numbers suggests itself as potentially a useful tool.[4] As long as the output functions F_o, the types in Γ, Δ, and the background condition B are first-order definable over the field of real numbers, as they are in our example, Tarski's decision procedure gives us a mechanical way to determine whether or not $\Gamma \vdash_B \Delta$.

Logic and Cognition

Much more speculatively, the proposal made here suggests a way out of the box that logic has been put into by some of its detractors. Within the recent cognitive-science literature, logic is often seen as irrevocably wed to what is perceived to be an outdated symbol-processing model of cognition. From there, it is but a short step to the conclusion that the study of logic is irrelevant for cognitive science. This step is often taken in spite of the fact that human reasoning is a cognitive activity and so must be part of cognitive science. Perhaps the use of state spaces might allow a marriage of logic with continuous methods like those used in dynamical systems and so provide a toehold for those who envision a distinctively different model of human reasoning (see Chapter 10 of Barwise and Perry (1983), for example). The (admittedly wild) idea is that the input and output of reasoning could be symbolic, at least sometimes, whereas reasoning itself might be better modeled by state-space equations, with the two linked together by means of something like infomorphisms.

[4] See Rabin (1974) for a brief exposition of Paul Cohen's improved proof of Tarski's theorem.

Lecture 20

Representation

The concepts of information and representation are, of course, closely related. Indeed, Jerry Fodor feels that they are so closely related as to justify the slogan "No information without representation." Though we do no go that far, we do think of the two as intimately connected, as should be clear from our account. In this lecture, we sketch the beginnings of a theory of representation within the framework presented in Part II. We have three motives for doing so. One is to suggest what we think such a theory might look like. The second is to explain some interesting recent work on inference by Shimojima. The third is to show that Shimojima's work has a natural setting in the theory presented here.[1]

20.1 Modeling Representational Systems

When we think of information flow involving humans, some sort of representational system is typically, if not always, involved: Spoken or written language, pictures, maps, diagrams, and the like are all examples of representations. So representations should fit into our general picture of information flow.

A theory of representation must be compatible with the fact that representation is not always veridical. People often misrepresent things, inadvertently or otherwise. For this reason, we model representation systems as certain special kinds of information systems where unsound logics can appear. We begin with our model of a representational system.

Definition 20.1.

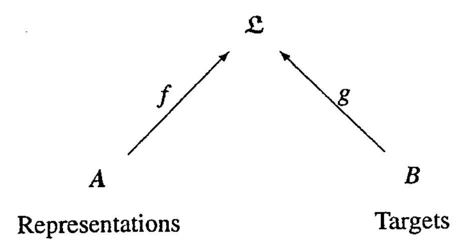
1. A *representation system* $\mathcal{R} = \langle \mathcal{C}, \mathfrak{L} \rangle$ consists of a binary channel $\mathcal{C} = \{f :$

[1] Representation is such a big and important topic that it surely deserves a book of its own. We hope one of our readers will write a book that further develops the ideas sketched here.

$A \rightleftarrows C, g : B \rightleftarrows C\}$, with one of the classifications designated as *source* (say A) and the other as *target*, together with a local logic $\mathfrak{L}$ on the core C of this channel.

2. The *representations* of $\mathcal{R}$ are the tokens of A. If $a \in \text{tok}(A)$ and $b \in \text{tok}(B)$, a is a *representation of* b, written $a \rightsquigarrow_{\mathcal{R}} b$, if a and b are connected by some $c \in C$. The token a is an *accurate representation of* b if a and b are connected by some normal token, that is, some $c \in N_{\mathfrak{L}}$.
3. A set of types Γ of the source classification *indicates* a type β of the target classification, written $\Gamma \Rightarrow_{\mathcal{R}} \beta$, if the translations of the types into the core gives us a constraint of the logic $\mathfrak{L}$, that is, if $f[\Gamma] \vdash_{\mathfrak{L}} g(\beta)$. The *content* of a token a is the set of all types indicated by its type set. The representation a *represents* b *as being of type* β if a represents b and β is in the content of a.

We can depict a representational system as follows:



The connections between the representations and their targets model the particular spatial-temporal process whereby the representation comes to represent what it does. The constraints of the logic model the various sort of constraints on this process. When dealing with conventional representation systems like writing or map making, the constraints will typically involve a wide variety of constraints of different types, from conventional to physically necessary, all the way to logic constraints. In this sketch, we simply group them all together.

With our definition, not every representation of $\mathcal{R}$ is a representation of some token in the target. (Novels are representations, but not necessarily of anything real.) Also note that a representation may be a representation of more than one token. For example, the picture of Ben taken on his third birthday represents him as he was then, but it also represents the Ben of today, two years later, though not as accurately as it represents the younger Ben.

Example 20.2. Let us see how we can view the practice of mapmaking as a representation system under this definition. The source classification A is a classification of maps (the tokens) by means of their syntactic types, that is,

types that classify maps according to what is printed or drawn on them. The target classification consists of regions classified by properties of and relations between the things in the regions. Thus a typical source type might be something roughly like "a school icon labeled 'Harmony' is next to road line labeled '2nd' " and a target type might be "Harmony School is located on 2nd street." The core of this classification models the actual practice of mapmaking. The connections are causal links between maps and what they are maps of. The types are ways of classifying these links. The logic on this classification models the understanding of users of these maps. Thus the constraints of the system represent the assumptions users of the maps make about the links between maps and what they are maps of. The normal tokens of the logic are the normal links, which must satisfy the constraints. Note, however, that there may be links that do not satisfy all of the constraints of the logic.

Assume the map a is a representation of Mt. Ateb. Suppose that looking at a we observe that $\Gamma \subseteq \text{typ}(a)$, that is, that a of every type in Γ. If $\Gamma \Rightarrow_{\mathcal{R}} \beta$, we would be justified in saying that the map a represents Mt. Ateb as being of type β, whether or not Mt. Ateb is of type β.

The next result shows that our definition of an accurate representation behaves properly.

Proposition 20.3. *If a is an accurate representation of b and a represents b as being of type β, then $b \vDash_B \beta$.*

Proof. Because a is an accurate representation of b, there is a normal connection c between a and b, so $f(c) = a$ and $g(c) = b$. Because a represents b as being of type β, $f[\text{typ}(a)] \vdash_{\mathfrak{L}} g(\beta)$. Because f is an infomorphism, $c \vDash_C f(\alpha)$ for each $\alpha \in \text{typ}(a)$. Because c is normal, c satisfies all constraints of the logic, so $c \vDash_C g(\beta)$. Because g is an infomorphism, $b \vDash_B \beta$. □

20.2 Imperfect Representations

In this section, we sketch our solution to a potentially troublesome fact about information and representation. Suppose we have a map of Mt. Ateb but the mountain has changed in small ways since the map was produced. Maybe a portion of a path has been obliterated by a landslide. There is clearly a sense in which the map is no longer accurate. After all, it represents the existence of a path where there is in fact none. Still, the map does carry a lot of valid information about the mountain. Any informational theory of representation should be compatible with this commonplace phenomenon.

Recall that our tokens in the target classifications are really regions at times, not just regions outside of time. Let us suppose that the original connection c_0 between the map and the mountain was normal, and connected the map a to m_0. Thus a was a representation of m_0. But when the path was destroyed, this changed m_0 to m_1 and, as a result, the connection c_0 between a and m_0 gave rise to a new connection c_1 between a and m_1. Thus a represents both m_0 and m_1. The new connection supports most of the constraints of the representation system but not all of them. In particular, it fails to support the constraint that says that a path icon at location l indicates a path at position p.

Whether or not a representation a is an *accurate* representation of b depends on the connections and the representation system in question; in particular, it depends on the local logic $\mathfrak{L}$ at the core of the system. It may be inaccurate with respect to a logic $\mathfrak{L}$ but accurate with respect to a slightly weaker logic $\mathfrak{L}_0 \sqsubseteq \mathfrak{L}$. The connection c between a and b is not normal with respect to $\mathfrak{L}$ but is normal with respect to a slightly weaker logic $\mathfrak{L}_0$. Relative to this logic, and the associated representation system, a is accurate.

20.3 Shimojima's Thesis

When we use representations to reason about some domain, the nature of the representations can greatly affect the form of the reasoning. The choice of a good representation system for the problem at hand can make the reasoning much easier. Conversely, the choice of a poor representation system can make the reasoning more difficult or even impossible.

There have been many attempts made to explain the various properties of different kinds of representation systems. One common intuition is that the better a fit there is between the representing domain and the represented domain, the better the representational system is. Atsushi Shimojima (1996) has used the basic notions of classification and constraint to give a rigorous formulation of this basic intuition, and has used it to investigate a wide range of representational phenomena. With various case studies, Shiomjima makes a strong case for the following:

Shimojima's Constraint Thesis

The inferential advantages and disadvantages between representation systems with a common target classification can be understood in terms of the kinds of constraints that the system projects from representations to targets.[2]

[2] Shimojima never goes quite so far as to call this a thesis. Rather, he calls it an hypothesis. We think his dissertation makes such a strong case, however, that we promote it to a thesis here.

In this section, we give a brief introduction to Shimojima's (1996) work after first setting it in the framework of representation systems as modeled in this chapter.

For a rigorous treatment of the thesis, Shimojima must first give an analysis of what it means for one constraint to project to another constraint.[3] Recall we have a representation system $\mathcal{R}$ with source classification $\boldsymbol{A}$ and target classification $\boldsymbol{B}$. For sets $\Gamma \subseteq \mathrm{typ}(\boldsymbol{A})$ and $\Theta \subseteq \mathrm{typ}(\boldsymbol{B})$, we say that $\mathcal{R}$ *projects* Γ *to* Θ, and write $\Gamma \Rightarrow^* \Theta$, if every $\beta \in \Theta$ is indicated by some $\alpha \in \Gamma$ and every $\alpha \in \Gamma$ indicates some $\beta \in \Theta$. Here are some simple properties of this relation. Recall that $a \models_{\wedge A} \Gamma$ means that $a \models_A \alpha$ for every $\alpha \in \Gamma$, and dually for disjunction. The following is easy to verify.

Proposition 20.4. *Let a be an accurate representation of b and assume that $\mathcal{R}$ projects Γ to Θ.*

1. *If $a \models_{\wedge A} \Gamma$, then $b \models_{\wedge B} \Theta$.*
2. *If $a \models_{\vee A} \Gamma$, then $b \models_{\vee B} \Theta$.*
3. *$\Gamma = \emptyset$ if and only if $\Theta = \emptyset$.*

Furthermore, if $\Gamma = \{\alpha\}$ and $\Theta = \{\beta\}$, then $\Gamma \Rightarrow^* \Theta$ if and only if $\alpha \Rightarrow_{\mathcal{R}} \beta$.

Definition 20.5. Let $S = \langle \Gamma, \Delta \rangle$ and $S' = \langle \Theta, \Psi \rangle$ be sequents of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively. We say that $\mathcal{R}$ *projects* S to S' if $\Gamma \Rightarrow^* \Theta$ and $\Delta \Rightarrow^* \Psi$. If S is a constraint of $\boldsymbol{A}$, then this relationship is said to be an instance of *constraint projection.*

Let a be an accurate representation of b and assume that $\mathcal{R}$ projects S to S'. One might be tempted to think that if a satisfies the sequent S, then b satisfies the sequent S'. A moment's thought shows that this is not in general the case. Indeed, part of Shimojima's argument in favor of his thesis comes from exploring cases where projected constraints are not in fact satisfied. We give a hint as to his argument by examining three cases of projection: where the consequent of S is empty, where it is a singleton, and where it has more than one element.

Free Rides

Shimojima defines a *free ride* to be any case of constraint projection $S \Rightarrow S'$ where both S and S' have a single type in the succedent of the sequent. That is,

[3] Actually, Shimojima did not use the framework of channel theory, but instead used the notions of signaling and indicating as primitives.

S is of the form $\langle\Gamma, \{\alpha\}\rangle$ and $S' = \langle\Theta, \{\beta\}\rangle$. In other words, a free ride occurs if $\Gamma \Rightarrow^* \Theta$, $\Gamma \vdash_A \alpha$, and $\alpha \models_{\mathcal{R}} \beta$.

Suppose we want to represent the information (or misinformation) Θ about some b by means of a representation a whose type set includes Γ. Any such representation will represent its target b as being of all the types in Θ. Note, however, that because $\Gamma \vdash_A \alpha$, a will also be of type α. But then because $\alpha \Rightarrow_{\mathcal{R}} \beta$, a will also represent b as being of type β. In other words, any attempt to represent b as satisfying all the types in Θ by means of a representation satisfying Γ will automatically represent b as being of type β. This can be a good or a bad thing.

Positive Free Rides

If $\Theta \vdash_B \beta$, then this is called a positive free ride, because the conclusion β automatically generated by creating a representation satisfying Γ is in fact warranted. One of the advantages of various forms of diagrammatic representations is that one gets various forms of positive free rides.

Example 20.6. Let us go back to the example of maps. The cartographer places the line l_1 representing Atwater Street above the line l_2 representing Second Street so the map represents Atwater as being north of Second. The cartographer similarly places the line l_3 representing University Street below l_2 so the map represents University as being south of Second. By virtue of placing l_1 above l_2 and l_3 below l_2, the cartographer also places l_1 above l_3, by virtue of a constraint S of $\boldsymbol{A}$. As a result, the map automatically represents Atwater as being north of University. This is a legitimate piece of information because the constraint S projects to a matching constraint S' on $\boldsymbol{B}$.

Negative Free Rides

If $\Theta \nvdash_A \beta$, then this is called a negative free ride, because the conclusion β generated automatically by creating a representation satisfying Γ is not warranted. The mismatch between distances between points on a planar map of a curved surface (like the earth) and the distances between the points they represent gives rise to examples of this.

Overdetermined Alternatives

By a case of *overdetermined alternatives* Shimojima means any case of constraint projection $S \Rightarrow S'$ where both S and S' have more than one type in the succedent, say $S = \langle\Gamma, \Delta\rangle$ and $S' = \langle\Theta, \Psi\rangle$, where Ψ has more than one element but where for each $\beta \in \Psi$, $\Theta \nvdash_B \beta$.

Suppose we were to represent b as satisfying all the types in Θ by means of a representation a satisfying all types in Γ. Because S is a constraint, a must satisfy some $\alpha \in \Delta$. But then by Proposition 20.4.2, there would be some $\beta \in \Psi$ such that a represents b as being of type β. But because $\Theta \nvdash_B \beta$, our representation would have represented b as being of some type that does not follow from the information Θ from which we started. That is, any representation a we choose that represents b as satisfying all the types in Θ by means of Γ has the unfortunate effect of representing b as satisfying some β that does not follow from Θ.

Example 20.7. Suppose the cartographer knows there is new high school being planned for south of town and that the school will be built quite soon. Thus he or she would like to indicate it on the map. However, if he places a school icon at any particular place on the map, it will indicate that the school is at a particular location. Although one of the locations is the right one, which one does not follow from the information Θ at the cartographer's command, that there will be a school south of town.

As with the case of free rides, there are two subcases to consider, depending on whether or not S' is a constraint of $\boldsymbol{B}$. If S' is not a constraint of $\boldsymbol{B}$, then overdetermined alternatives are a potential source of serious error when using the representation system. However, if S' is a constraint, there is a way to get around the problem.

Cases Exhaustive

If S' is a constraint, that is, if $\Theta \vdash_B \Psi$, then one can work around the difficulty by using multiple representations in a disjunctive fashion. To simplify the discussion, let us suppose that $\Delta = \{\alpha_1, \alpha_2\}$ has two types. Suppose that we find representations a_1 and a_2 of b satisfying all types in Γ with $a_1 \models_A \alpha_1$ and $a_2 \models_A \alpha_2$. These will be called a set of *exhaustive alternatives* for $S \Rightarrow^* S'$. Because each type in Ψ is indicated by at least one of these two, and because both possibilities in Δ are covered, between them they legitimately represent b as satisfying at least one type in Δ without representing b as being of any particular type in Δ.

Example 20.8. Cases exhaustive is one of the main methods of reasoning with simple diagrammatic systems, where one has to break a piece of reasoning into a number of distinct cases. Suppose, for example, we were given a representation

representing a seating arrangement involving Greg, Mary, Bill, Ann, and Toni and we were given the information that Ann was seated next to Mary. If we place an A in the first empty box, it will represent one alternative and represent Ann as sitting next to Greg, in the second a different alternative will represent Ann as sitting next to Bill. We need to break into two exhaustive cases, from which it will follow that Ann is next to either Greg or Bill. For further discussion of the importance of this rule, see the discussion of the Hyperproof system's rule of Cases Exhaustive in Barwise and Etchemendy (1995).

The validity of the method of cases exhaustive depends crucially on the validity of the projected constraint, as Shimojima's analysis makes clear.

Autoconsistent Systems

One of the advantages of certain kinds of representations is that it is difficult, sometimes impossible, to use them to represent inconsistent information. This means that one can demonstrate consistency of some purported information simply by producing a representation of it in such a system. For example, a drawing might be used to show that a certain placement of furniture in the room is possible.

The system $\mathcal{R}$ is said to be *autoconsistent* if whenever $\mathcal{R}$ projects a sequent S of the form $\langle \Gamma, \emptyset \rangle$ to some sequent $S' = \langle \Theta, \emptyset \rangle$, if S' is a constraint, so is S.

Example 20.9. The representation system employed in Hyperproof is autoconsistent. The standard Euler circle system is also autoconsistent.

Proposition 20.10. *Suppose $\mathcal{R}$ is autoconsistent. For every $a \in$ tok(A), the content of a holds for some token $b \in$ tok($\boldsymbol{B}$).*

Proof. Let $\Gamma = \text{typ}(a)$ and Θ be the content of a. By definition of content, $\Gamma \Rightarrow^* \Theta$. If Θ holds for no token of $\boldsymbol{B}$, then $\Theta \vdash_{\boldsymbol{B}}$. By autoconsistency, $\Gamma \vdash_{\boldsymbol{A}}$, which contradicts the fact that a satisfies every type in Γ. □

Shimojima (1996) goes into these ideas in much greater depth and with many more examples. He also has an interesting discussion of the roles that different types of constraints, say nomic versus conventional, play in different kinds of representational systems. We recommend it to the reader.

Lecture 21

Quantum Logic

Quantum theory is well known for having a nonclassical and somewhat peculiar logic. One approach to trying to make this logic comprehensible is by means of what is called a "manual of experiments." This is the approach taken, for example, in the book *An Introduction to Hilbert Space and Quantum Logic* (Cohen, 1989), which we use as our main reference. A manual is thought of as describing a variety of different tests of a system, but where the making of one test may, for some reason, preclude the making of some other test of the same system at the same time. Still, it can happen that from performing one kind of experiment we may get information about what the outcome would have been had we performed another kind of experiment.

Example 21.1. Here is an example taken from 17). Imagine a firefly trapped inside a box, and two kinds of experiments, FRONT and SIDE, one can perform when observing this box at a given instant. In a FRONT experiment one looks in the box from the front, whereas in SIDE one looks from the right-hand side. In FRONT there are three possible outcomes: one may see the firefly light lit up on the right (R), lit up on the left (L), or not lit up at all (N). Similarly, in SIDE there are three possible outcomes: one may see the firefly light at the front (F), at the back (B), or not at all (N). Suppose an observer examines the box using FRONT and observes an outcome of L. Then the firefly was lit, so if the observer had performed the experiment SIDE, one would have obtained an outcome in the set {F, B}. Thus we have a constraint L $\vdash$ F, B, and associated information flow.

We want to see how this kind of information flow fits into the picture presented in this book. We also want to see how well the framework of this book fits with that of quantum logic.

21.1 The Theory of a Manual

Definition 21.2. A *manual of experiments* is an indexed family $M = \{\Omega_E\}_{E\in\mathcal{E}}$ of sets. Each $E \in \mathcal{E}$ is called an experiment (or kind of experiment). The set Ω_E is called the set of *possible outcomes* of E.

Cohen calls this a "quasi-manual." What he calls a manual we will call a standard manual (Definition 21.17). Each set Ω_E can be thought of as the set of states of a state-space $\boldsymbol{S}_E$. To make things canonical, we take $\boldsymbol{S}_E$ to be the ideal state space with Ω_E as states. Thus Ω_E is the set of types and the set of tokens of $\boldsymbol{S}_E$, and each token is its own type. In this way, each E gives rise to a theory, namely, the $\text{Th}(\boldsymbol{S}_E)$. The types of this theory are arbitrary subsets of Ω_E and $\Gamma \vdash_{\text{Th}(\boldsymbol{S}_E)} \Delta$ if and only if $\bigcap \Gamma \subseteq \bigcup \Delta$.

We want to associate a sound and complete local logic $\text{Log}(M)$ with any manual M. We start with the manual's theory $\text{Th}(M)$, which has an obvious natural definition. From that, we work toward a definition of a classification $\text{Cla}(M)$ associated with M and then show that $\text{Th}(M)$ is the theory of the classification $\text{Cla}(M)$.

Definition 21.3. For any manual $M = \{\Omega_E\}_{E\in\mathcal{E}}$, $\text{Th}(M)$ is the theory whose set of types is

$$\Sigma_M = \bigcup_{E\in\mathcal{E}} \text{pow}\,\Omega_E$$

and whose consequence relation is the smallest regular consequence relation containing the consequence relation $\vdash_{\boldsymbol{S}_E}$ of $\text{Log}(\boldsymbol{S}_E)$, for each $E \in \mathcal{E}$.

Example 21.4. In the firefly manual M, the set $\mathcal{E}$ of experiments is {FRONT, SIDE}, $\Omega_{\text{FRONT}} = \{\text{L}, \text{R}, \text{N}\}$ and $\Omega_{\text{SIDE}} = \{\text{F}, \text{B}, \text{N}\}$. Σ_M consists of all subsets of either of these sets. It does not contain a set like $\{\text{F}, \text{L}\}$. Because $\{\text{L}\}, \{\text{N}\} \vdash_{\boldsymbol{S}_{\text{FRONT}}}$ and $\vdash_{\boldsymbol{S}_{\text{SIDE}}} \{\text{N}\}, \{\text{F}\}, \{\text{B}\}$, both sequents hold in $\text{Th}(M)$. Hence by Cut and Weakening, we obtain the constraint

$$\{\text{L}\} \vdash_{\text{Th}(M)} \{\text{F}\}, \{\text{B}\}$$

as desired.

We want to show that the theory $\text{Th}(M)$ of a manual M is the complete theory of a classification naturally associated with M. To motivate the choice of tokens, we first present some definitions from quantum logic.

Operationally Perspective Types

Let M be a manual and let Σ_M be the associated set of types. A pair of types α, β are said to be *comeasurable* if there is an experiment E of the manual so that $\alpha \cup \beta \subseteq \Omega_E$. That is, they are comeasurable if you can test for them both with a single experiment E. In the firefly example, {L} and {R} are comeasurable but not consistent. On the other hand, {L, N} and {B, N} are not comeasurable, even though they could both hold in a given situation.

Types α and β are *orthogonal*, written $\alpha \perp \beta$, if they are comeasurable and disjoint. In the firefly example, $\{\text{L}, \text{R}\} \perp \{\text{N}\}$. Types α and β are *orthogonal compliments*, written α oc β, provided they are orthogonal and their union is an Ω_E, that is, if the two events partition the set Ω_E of possible outcomes of some experiment $E \in \mathcal{E}$. Intuitively, this means that exactly one of the two events would be seen to hold were we to perform an experiment of type E. (In the firefly example, {L, R} and {N} are orthogonal compliments.)

Types α_1 and α_2 are *operationally perspective*, α_1 op α_2, if there is a type β such that α_1 oc β and α_2 oc β. Intuitively, operationally perspective types are equivalent, because they have a common compliment. (In the firefly example, {L, R} op {F, B} because they are both orthogonal complements of {N}.)

In the approach to quantum logic using manuals, it is assumed that operationally perspective types are equivalent. We adopt this assumption and see where it leads us in our search for a classification to go along with the theory $\text{Th}(M)$. First, though, note the following result.

Proposition 21.5. *Let M be a manual and let* $\text{Th}(M) = \langle\Sigma_M, \vdash\rangle$ *be its theory. For all* $\alpha, \beta \in \Sigma_M$*:*

1. *If* $\alpha \perp \beta$, *then* $\alpha, \beta \vdash$.
2. *If* α oc β, *then* $\alpha, \beta \vdash$ *and* $\vdash \alpha, \beta$.
3. *If* α op β, *then* $\alpha \vdash \beta$ *and* $\beta \vdash \alpha$.

Proof. To prove (1), assume $\alpha \perp \beta$. Then α, β are disjoint but are both subsets of Ω_E for some E. But then $\alpha, \beta \vdash_{\boldsymbol{S}_E}$ so $\alpha, \beta \vdash$. The proof of (2) is similar so we prove (3). Because op is a symmetric relation, it suffices to prove one of the two conclusions. Assume that α op β. Let θ be such that α oc θ and θ oc β. By (2), $\alpha, \theta \vdash$ and $\vdash \theta, \beta$. But then by Finite Cut, we have $\alpha \vdash \beta$. □

Let

$$A = \sum_{E\in\mathcal{E}} \text{Evt}(\boldsymbol{S}_E)$$

and let

$$\mathfrak{L} = \sum_{E \in \mathcal{E}} \text{Log}(S_E),$$

the corresponding sum of logics. Being a sum of sound and complete logics, $\mathfrak{L}$ is a sound and complete logic on A. The tokens are functions c that assign to each E some element $c_E \in \Omega_E$. We can think of such a token as representing the possible outcomes of performing the different experiments in $\mathcal{E}$: c_E is the outcome one would get were one to perform an experiment of kind E. The types are the disjoint union of the types of the various state spaces.

Recall that typ(A) is the *disjoint* union of pow Ω_E, for $E \in \mathcal{E}$. But if $\alpha \in$ pow Ω_E and $\beta \in$ pow Ω_F, they are already distinct objects unless they are identical, in which case they are subsets of $\Omega_E \cap \Omega_F$. But these are exactly the types that give rise to op relationships; that is, if $\alpha = \beta$, then $\Omega_E - \alpha$ is operationally perspective to $\Omega_F - \beta$. So in this case we do not really want to make α and β distinct.

We need, then, to undo the disjointness caused by taking the sum. This is easily accomplished by taking the appropriate (dual) quotient. Define a dual invariant $J = \langle C, R \rangle$ on A as follows. The relation R relates exactly those types that are copies of the same set. The set C consists of those tokens of A that respect the relation R.

Lemma 21.6. *A token of A respects R if and only if for all types $E, F \in \mathcal{E}$, if $c_E \in \Omega_F$ then $c_E = c_F$.*

With this in mind, we make the following important definition.

Definition 21.7. An *idealized token* of M is an element $c \in \text{tok}(A)$ such that for all experiments $E, F \in \mathcal{E}$, if $c_E \in \Omega_F$ then $c_E = c_F$.

Proposition 21.8. *The logic $\mathfrak{L}/J$ is a sound and complete logic on the classification A/J.*

Proof. This is an immediate consequence of Corollary 12.34. □

Classification A/J has the idealized tokens for tokens. Its types are (up to isomorphism) just the types Σ_M of Th(M). This suggests that Th(M) is the theory of the sound and complete logic with idealized tokens as tokens and Σ_M as types.

Definition 21.9. Given a manual M, the *classification* Cla(M) of M is the classification whose types are the set Σ_M, whose tokens are the M-idealized tokens, and whose classification relation is given by $c \vDash \alpha$ if and only if $c_E \in \alpha$ for some (hence every) E such that $\alpha \subseteq \Omega_E$.

Justification. The classification relation is well defined by the definition of idealized tokens. □

Theorem 21.10. *For any manual M, the theory* Th(M) *is the sound and complete theory of* Cla(M). *Thus, for any sequent $\langle \Gamma, \Delta \rangle$ of Σ_M, $\Gamma \vdash_{\text{Th}(M)} \Delta$ if and only if every idealized token satisfies $\langle \Gamma, \Delta \rangle$.*

Proof. Let $f : \text{Cla}(M) \rightleftarrows A/J$ be the obvious token-identical infomorphism. It is clear that this is also a theory infomorphism taking Th(M) into the theory of $\mathfrak{L}/J$. Thus $\text{Th}(M) \sqsubseteq f^{-1}[\text{Th}(\mathfrak{L}/J)]$. To prove they are equal, we need only check that if $\langle \Gamma, \Delta \rangle$ is a consistent partition of Th(M), then there is an idealized token that is a counterexample to it. Define c_E to be the unique $\alpha \in \Omega_E$ such that $\{\alpha\} \in \Gamma$. It is easy to verify that if $c_E \in \Omega_F$, then $c_E = c_F$, because otherwise the restriction of the sequent to Ω_F would not be consistent in Log(S_F). □

This theorem shows that the theory Th(M) is sound as long as it is used with tokens that give rise to idealized tokens. It will be complete as long as every idealized token corresponds to a real token of the classification.

The theory Th(M) does not in general have a disjunction or a conjunction, as the firefly example shows. It does have a negation, however.

Corollary 21.11. *Suppose that α is both an E-type and an F-type. Then $\Omega_E - \alpha \vdash_{\text{Th}(M)} \Omega_F - \alpha$ and $\Omega_F - \alpha \vdash_{\text{Th}(M)} \Omega_E - \alpha$.*

Proof. See Exercise 21.2. □

We can use this corollary to show that the theory Th(M) has a negation. Define $\neg\alpha$ to be $\Omega_E - \alpha$ for some E such that α is an E-type. By 21.11, it does not matter which E we choose, they will all be equivalent. The theory obeys the usual rules for negation. This negation is usually called the "orthocomplement" in quantum logic.

Finally, we note the following disquieting possibility. Recall that a logic $\mathfrak{L}$ is coherent if not every sequent is a constraint, which is equivalent to saying that the empty sequent is not a constraint. If a logic has even a single normal token, then it is coherent.

Corollary 21.12. *Let M be a manual. There are idealized tokens for M if and only if the theory* Th(M) *is coherent.*

This corollary implies that if there are no idealized tokens for M, then the least logic obtained from putting together the state-space logics from the individual experiments is incoherent! We will see that there are manuals of use in quantum mechanics where this happens.

21.2 A Comparison with Quantum Logic

We have used our general framework to define a notion of consequence between the types in a manual and established an understanding of this logic in terms of idealized tokens. The usual notion of consequence from quantum logic proceeds rather differently. We review this development here and then compare it with that from above.

Definition 21.13. Given a manual M, α_1 is said to *M-entail* α_2, written $\alpha_1 \leq_M \alpha_2$, if and only if there is some β such that $\alpha_1 \perp \beta$ and $\alpha_1 \cup \beta$ op α_2.

Proposition 21.14. *Let M be a manual. If $\alpha_1 \leq_M \alpha_2$, then $\alpha_1 \vdash_{\mathrm{Th}(M)} \alpha_2$.*

Proof. Assume $\alpha_1 \perp \beta$ and $\alpha_1 \cup \beta$ op α_2. Because $\alpha_1 \perp \beta$, $\alpha_1 \cup \beta \subseteq \Omega_E$, for some $E \in \mathcal{E}$, so we have $\alpha_1 \vdash_{\mathrm{Log}(S_E)} \alpha_1 \cup \beta$, and hence $\alpha_1 \vdash_{\mathrm{Th}(M)} \alpha_1 \cup \beta$. But because $\alpha_1 \cup \beta$ op α_2, we have $\alpha_1 \cup \beta \vdash_{\mathrm{Th}(M)} \alpha_2$, by Proposition 21.5.3. We obtain $\alpha_1 \vdash_{\mathrm{Th}(M)} \alpha_2$ by Cut. □

Two-Experiment Manuals

The following result shows that in the case of a manual with two experiments, the notion of entailment from quantum logic coincides with that in the logic of the manual. We will turn to the more general question in the next section.

Theorem 21.15. *Suppose the manual M has only two experiments, E and F. If $\alpha \subseteq \Omega_E$ and $\beta \subseteq \Omega_F$ are types in α, then $\alpha \leq_M \beta$ if and only if $\alpha \vdash \beta$.*

Proof. We will show the following are equivalent:

1. $\alpha \leq_M \beta$;
2. $\alpha \vdash \beta$;
3. either $\alpha \subseteq \beta$ or $\alpha \cap \Omega_F \subseteq \beta$ and $\Omega_F \subseteq \beta \cup \Omega_E$.

We have already established (1) implies (2), so we need only show (2) implies (3) and (3) implies (1). Let us first establish the latter implication. Because (3) is a disjunction, there are two cases to consider. If $\alpha \subseteq \beta \subseteq \Omega_F$, then let $\theta = \beta - \alpha$ and $\gamma = \Omega_F - \beta$. Then $\alpha \perp \theta$ and $\alpha \cup \theta$ op β because $\alpha \cup \theta = \beta$ and β oc γ. The other case to consider is when $\alpha \cap \Omega_F \subseteq \beta$ and $\Omega_F \subseteq \beta \cup \Omega_E$. In this case, let $\gamma = \Omega_F - \beta$ so that γ is an orthogonal complement of β. Notice that $\gamma \subseteq \Omega_E$ because $\Omega_F \subseteq \beta \cup \Omega_E$. Now let θ consist of those states in Ω_E that are in neither α nor γ. Clearly, $\alpha \perp \theta$. We want to show that $\alpha \cup \theta$ op β by showing that γ is an orthogonal complement to $\alpha \cup \theta$ (as well as to β, as we saw). Notice that $\gamma \cap \alpha = \emptyset$ because $\alpha \cap \Omega_F \subseteq \beta$. Hence $\Omega_E - \gamma = \alpha \cup \theta$ as desired.

Now assume (2). We show that $\alpha \subseteq \beta$ or that $\alpha \cap \Omega_F \subseteq \beta$ and $\Omega_F \subseteq \beta \cup \Omega_E$. It is convenient to split the proof into two cases, depending on whether or not $\Omega_F \subseteq \beta \cup \Omega_E$.

Case 1. $\Omega_F \subseteq \beta \cup \Omega_E$. We prove that $\alpha \cap \Omega_F \subseteq \beta$ (and hence (3)). Suppose this is not the case. Let $\sigma \in \alpha \cap \Omega_F$, $\sigma \notin \beta$. Notice σ is an outcome in both E and F, because $\sigma \in \alpha \subseteq \Omega_E$. Hence there is an idealized token c that assigns σ to both E and F. But this token is of type α and not of type β, which contradicts (2).

Case 2. $\Omega_F \not\subseteq \beta \cup \Omega_E$. In this case we want to show that $\alpha \subseteq \beta$. Suppose this were not the case. Let $\sigma \in \alpha - \beta$ and let $\tau \in \Omega_F - (\beta \cup \Omega_E)$. If $\sigma \in \Omega_F$, then consider the idealized token c that assigns σ to both E and F. This shows that $\alpha \nvdash_{\mathrm{Th}(M)} \beta$. If $\sigma \notin \Omega_F$, then consider the idealized token c that assigns σ to E and τ to F. This idealized token shows that $\alpha \nvdash_{\mathrm{Th}(M)} \beta$. □

More Than Two Experiments

In Theorem 21.15, the assumption that the manual have only two experiments is crucial, as the following example shows.

Example 21.16. Here is a simple example with four experiments, each with two possible outcomes, where $\leq_M$ is weaker than $\vdash$. Let E have outcomes $\{1, 2\}$, F have outcomes $\{2, 3\}$, G have outcomes $\{3, 4\}$, and H have outcomes $\{4, 5\}$. Then $\{1\} \vdash \{5\}$ because the only idealized token that has $c_E = 1$ must also have $c_F = 3$ so $c_G = 3$ so $c_H = 5$. But it is also clear that $\{1\} \not\leq_M \{5\}$. This example also shows that the ordering $\leq_M$ is not even transitive, because $\{1\} \leq_M \{3\}$ and $\{3\} \leq_M \{5\}$.

This example shows that, in general, the quantum logic relation $\alpha \leq_M \beta$ is too weak to be reasonable, it is not even transitive. Hence it does not satisfy even Finite Cut. One approach would be to close the logic under such rules. However, a different approach is taken in the quantum logic literature, where it is standard practice to postulate additional structure on the manual.

Definition 21.17. A manual M is *standard* if it satisfies the following conditions:

1. If $E, F \in \mathcal{E}$, and $\Omega_E \subseteq \Omega_F$, then $\Omega_E = \Omega_F$.
2. If α_1 oc α_2, α_2 oc α_3, and α_3 oc α_4, then $\alpha_1 \perp \alpha_4$.
3. If $x \perp y$, $y \perp z$, and $z \perp x$, then the set $\{x, y, z\}$ is a type, that is, it is a subset of Ω_E for some $E \in \mathcal{E}$.

The manual of Example 21.16 does not satisfy condition (2) so it is not standard.

The first condition seems reasonable enough. After all, why have F in your manual if in fact some of its outcomes are impossible and can be eliminated by E?[1]

Why (2) is reasonable from an intuitive point of view is not at all obvious. It requires that we be able to combine experiments in a certain way. In particular, it rules out a manual of the kind described in Example 21.16, a manual that seems perfectly reasonable from an intuitive point of view. Here is what Cohen (1989, p. 23) has to say by way of justification. Read $\alpha_1, \alpha_2, \alpha_3$, and α_4 for A, B, C, and D, respectively.

> Suppose we test for event A by performing experiment E, and A occurs. Then we know that B did not occur. Thus, if we had performed experiment F [an experiment with both B and C as events] then C would have occurred; so, if we had performed experiment G [an experiment with both C and D as events] then D would *not* have occurred. In summary, if we test for A, and A occurs, then testing for D would result in D not occurring. [The reverse is then observed.] Hence A and D are events that bear a special relationship to each other through E, F, and G, and it is not unnatural to require that there be a single experiment H that contains both A and D so that $A \perp D$ in M.

The first step of the argument clearly assumes that if A op C and A occurred under E, then C would have occurred under F, that is, that A and C are equivalent. It is this condition that led to our notion of an idealized token. The assumption is basically that the outcome of an experiment is completely determined by the state of the system, not other factors of the experimental event itself. For a case where this does not hold, consider the following example.

[1] Of course it could be that E is much more expensive in time or effort than F. That is why we have book reviews.

Example 21.18. A coin c stands on its edge on a table. We can perform two experiments on this coin. We can pick it up and flip it, with the outcomes of heads or tails: $\Omega_{\text{FLIP}} = \{\text{H}, \text{T}\}$. Or we can bang the tabletop to see which way the coin falls: $\Omega_{\text{BANG}} = \{\text{H}, \text{T}\}$. (With a U.S. penny, the probability of an H in the second experiment is close to 1.) Notice that this is a standard manual. Let $A = \{\text{T}\} = C$ and $B = \{\text{H}\}$. Thus A op C, but clearly there is no reason to suppose that just because A occurred in an experiment of kind FLIP, that C would have occurred under BANG. So even though we have a standard manual M, its theory is completely inappropriate for understanding information flow with regard to this system.

This example shows that if the outcome of an experimental event is partially determined by the event itself, rather than being completely determined by the state of the system, then op types do not bear a special relationship to one another. To put it the other way around, making operationally perspective types equivalent presupposes that the instantaneous state of the system can be adequately modeled by an idealized token.

A secondary point is that even if the physicist is in a context where the argument up to the final step ("...it is not unnatural to require...") seems justified, that is, where the system can be modeled by an idealized token, the final step seems unmotivated. It seems likely that the real motivation is that the condition is needed to guarantee the transitivity of the entailment relation $\leq_M$, something we can achieve more gracefully by means of regular theories.

Condition (3), the so-called "ortho-coherence condition," asserts that given any three pairwise incompatible mutual outcomes, each of which can be tested against the others, there is a single experiment that can decide among all three. This seems even less natural than the second condition. (Cohen remarks that this condition is not always assumed.)

These requirements on standard manuals, unnatural as they may appear in general, do hold in the manual naturally associated with a Hilbert space.

Given a Hilbert space H, an *orthonormal base* of H is a maximal orthogonal set E of unit vectors. An orthonormal basis E is the set of eigenstates of a commuting set of Hermitian operators where these operators are used to model experiments. We can think of the total state of a particle as modeled by a unit vector $\xi \in H$. Given an orthonormal basis E, $\xi = \sum_{v \in E} r_v v$ for some choice of scalars whose sum is 1. Performing the experiment modeled by E yields one of these vectors $v_\xi \in E$; intuitively, the outcome depends on the weights: the larger the weight, the more likely that vector is the outcome.

Definition 21.19. Let H be a Hilbert space. The *frame manual of* H, written $\mathcal{F}(H)$, is the manual defined as follows. The experiments E of $\mathcal{F}(H)$ consist of the maximal orthonormal subsets of H. The possible outcomes of E consist of the elements of E.

Proposition 21.20. *For any separable Hilbert space (which includes all finite-dimensional Hilbert spaces), the frame manual $\mathcal{F}(H)$ is standard.*

Proof. This is a standard result; see Example 3.4A of 17). □

The natural question to ask, in view of this result, is whether the consequence relation of $\mathrm{Th}(M)$ agrees with the entailment relation $\leq_M$ for standard manuals M. That is, can Theorem 21.15 be extended to arbitrary standard manuals? A very strong negative answer to this question was obtained by Kenneth Kunen. Michael Dickson observed that this is equivalent to the famous Kochen–Specker Theorem (29).

Theorem 21.21. *If H is a (real or complex) Hilbert space of dimension at least three, then the frame manual $\mathcal{F}(H)$ has no idealized tokens. In other words, for any function g such that $g(E) \in E$ for each orthonormal base E, there are two orthonormal bases E, F with $g(E) \in F$ but $g(E) \neq g(F)$.*

Proof. Because the theorem for complex Hilbert spaces easily follows from the theorem for real Hilbert spaces, we assume that H is a real Hilbert space. Suppose we had such an idealized token g. Let $\mathcal{E}$ be the set of all $g(E)$ such that E is an orthonormal base (the set of chosen vectors). Note that no two vectors in $\mathcal{E}$ are orthogonal and $E \cap \mathcal{E}$ is a singleton for each orthonormal base E.

Let $\angle(\vec{v}, \vec{w})$ denote the angle between two nonzero vectors, $\vec{v}, \vec{w}$. Because the unit sphere is connected, let $\vec{v}_1$ and $\vec{w}_1$ be unit vectors such that $\vec{w}_1 \in \mathcal{E}$ and $\vec{v}_1 \notin \mathcal{E}$ and $\angle(\vec{v}_1, \vec{w}_1) \leq 1°$. Now, $\mathcal{E}$ contains some, but not all, unit vectors in the complementary space $\vec{v}_1^{\perp}$, so again by connectedness, choose unit vectors $\vec{v}_2$ and $\vec{w}_2$, both orthogonal to v_1, such that $\vec{w}_2 \in \mathcal{E}$ and $\vec{v}_2 \notin \mathcal{E}$ and $\angle(\vec{v}_2, \vec{w}_2) \leq 1°$. Finally, because $\vec{v}_1, \vec{v}_2$ are orthogonal and not in $\mathcal{E}$, there is a unit vector $\vec{w}_3$ orthogonal to them both and in $\mathcal{E}$. So the $\vec{w}_i$ are all in $\mathcal{E}$ ($i = 1, 2, 3$), and $89° \leq \angle(\vec{w}_i, \vec{w}_j) \leq 91°$ whenever $i \neq j$. □

Lemma 21.22. *If H is a real Hilbert space, the $\vec{w}_i$ are all nonzero vectors ($i = 1, 2, 3$), and $89° \leq \angle(\vec{w}_i, \vec{w}_j) \leq 91°$ whenever $i \neq j$, then there is an orthonormal base E for H such that each vector in E is orthogonal to at least one of the $\vec{w}_i$.*

Assuming the lemma, we have a contradiction, because $g(E)$ and some $\vec{w}_i$ will yield two orthogonal vectors in $\mathcal{E}$.

To prove the lemma, note that it is sufficient to prove the lemma for three-dimensional Euclidian space (whence E will have size three), because in the general case, we can produce three orthogonal vectors in the linear span of the $\vec{w}_i$ and extend them arbitrarily to an orthonormal base. Now E will be $\{\vec{p}_1, \vec{p}_2, \vec{p}_3\}$.

Before obtaining E, we fix an orthonormal base, $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ "close to" to $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ as follows. Choose $\vec{e}_1$ parallel to $\vec{w}_1$. Obtain $\vec{e}_2 \perp \vec{e}_1$ in the $(\vec{w}_1, \vec{w}_2)$ plane by rotating $\vec{w}_2$ by $\leq 1°$, so $\angle(\vec{w}_2, \vec{e}_2) \leq 1°$. Then, obtain $\vec{e}_3$ by rotating $\vec{w}_3$ by $\leq 3°$.

Because the lengths of the $\vec{w}_i$ are irrelevant here, we may assume that

$$\begin{aligned}
\vec{w}_1 &= \vec{e}_1, \\
\vec{w}_2 &= a\vec{e}_1 + \vec{e}_2, \\
\vec{w}_3 &= b\vec{e}_1 + c\vec{e}_2 + \vec{e}_3,
\end{aligned}$$

where $|a|, |b|, |c| \leq \tan(3°) \leq 0.1$. Let $E = \{\vec{p}_1, \vec{p}_2, \vec{p}_3\}$, where

$$\begin{aligned}
\vec{p}_1 &= 0\,\vec{e}_1 + \cos(\theta)\,\vec{e}_2 + \sin(\theta)\,\vec{e}_3, \\
\vec{p}_2 &= \sin(\varphi)\,\vec{e}_1 - \cos(\varphi)\sin(\theta)\,\vec{e}_2 + \cos(\varphi)\cos(\theta)\,\vec{e}_3, \\
\vec{p}_3 &= \cos(\varphi)\,\vec{e}_1 + \sin(\varphi)\sin(\theta)\,\vec{e}_2 - \sin(\varphi)\cos(\theta)\,\vec{e}_3,
\end{aligned}$$

where θ and φ will be determined so that each $\vec{p}_i \perp \vec{w}_i$. Note that if $a = b = c = 0$, then we may take $\theta = \varphi = 0$, whence $\vec{p}_1 = \vec{e}_2$, $\vec{p}_2 = \vec{e}_3$ and $\vec{p}_3 = \vec{e}_1$. In the general case, the fact that a, b, c are small means that we may find small θ, φ that work.

For any values of θ, φ, our E is an orthonormal base and $\vec{p}_1 \perp \vec{w}_1$, so we must choose θ, φ to satisfy

$$a\sin(\varphi) - \cos(\varphi)\sin(\theta) = 0$$

$$b\cos(\varphi) + c\,\sin(\varphi)\sin(\theta) - \sin(\varphi)\cos(\theta) = 0$$

If $a = 0$, set $\theta = 0$ and choose φ so that $\tan(\varphi) = b$. If $a \neq 0$, choose θ so that $-45° \leq \theta \leq 45°$ and $f(\theta) = ab + c\,\sin^2(\theta) - \sin(\theta)\cos(\theta) = 0$; this is possible because $f(-45°) > 0 > f(45°)$. Then choose φ such that $a\,\sin(\varphi) - \cos(\varphi)\sin(\theta) = 0$ (i.e., $\tan(\varphi) = \sin(\theta)/a$). So, $0 = f(\theta)\cos(\varphi)/a = b\,\cos(\varphi) + c\,\sin(\varphi)\sin(\theta) - \sin(\varphi)\cos(\theta)$. For small a, b, c, we have (in radians) $\theta \approx ab$ and $\varphi \approx b$.

Corollary 21.23. *If H is a Hilbert space of dimension ≥ 3, then the local logic* $\text{Th}(\mathcal{F}(H))$ *of the frame manual is incoherent.*

Proof. This is an immediate consequence of the theorem and Corollary 21.12. In particular, $\alpha \vdash \beta$ for all α, β whereas there are α, β such that $\alpha \not\leq_M \beta$. □

This result shows that the Hilbert-space model of quantum mechanics is, in general, incompatible with the view of experiments as represented in our notion of the theory of a manual. Notice that the conflict has nothing at all to do with the meaning of any logical connectives. It has instead to do with two interacting decisions: the decision to have operationally perspective types be equivalent, and the decision to look at only regular theories, thereby building in the principle of Global Cut.

For lack of a better name, let us call a system *deterministic* with respect to a manual M if a token of the system partitions each set of types into those types the token is of and those types it is not of, and if which way a type common to more than one experiment falls is independent of the experiment performed. If we reexamine the proof of soundness of Global Cut in the present context, we see that it depends strongly on the assumption that the system is deterministic in this sense. Our example of the two-experiment manual involving a penny should make this clear. Using this terminology, we see that we cannot think of the set of all orthonormal bases of a Hilbert space (of dimension $\geq$3) as a manual for any deterministic system at all. Hence if the manual of a Hilbert space is a good model of the space of states of fundamental particles, then these particles are not deterministic.

Example 21.24. Recall the manual M for flipping and banging pennies. We can think of a penny c as having a state $\xi(c)$, something that characterizes the distribution of copper in the penny. If we perform either experiment, FLIP or BANG, the outcome will be strongly affected by $\xi(c)$. But it will also be effected by many other factors, like how hard we flip the coin, or at what rate it spins, or how hard we bang the table. These experiments can give us evidence about the state $\xi(c)$ but the outcome of these experiments neither determines nor is completely determined by $\xi(c)$.

Conclusion

It seems like the notion of operationally perspective types is basic to the project of using manuals to give an intuitive account of the logic of quantum mechanics. But it also seems that this notion presupposes that the system under investigation is deterministic in the sense defined above: that its total state can be modeled by an idealized token. For deterministic systems, we have characterized the resulting logic in two different ways, and everything seems quite sensible. But we have also seen that the family of orthonormal bases of a Hilbert space (of dimension $\geq$3) cannot be the manual of a deterministic system.

Exercises

21.1. What are the idealized tokens of the firefly manual and the penny manual?

21.2. Prove Corollary 21.11.

Answers to Selected Exercises

Lecture 4

4.1. The three classifications are discussed in turn.

- **4.4.** In this example typ(a) = the sets of which a is a member, tok(α) $= \alpha$, extensional and separated.
- **4.5.** In this example typ(a) $= \{f(a)\}$, tok(b) $= f^{-1}\{b\}$, extensional if f is surjective (or almost surjective: $\boldsymbol{B} - \mathrm{rng}(f)$ is a singleton), separated if f is injective.
- **4.6.** Finally, typ(M) = truths of L in M, tok(φ) = models making φ true. This is never extensional only because $\varphi \wedge \varphi$ is coextensive with but not identical to φ and never separated because isomorphic structures are indistinguishable.

4.2. For each type α of $\boldsymbol{A}$, $f(b) \models_{\boldsymbol{A}} \alpha$ if and only if $b \models_{\boldsymbol{B}} f(\alpha)$, by the infomorphism condition, if and only if $b' \models_{\boldsymbol{B}} f(\alpha)$, by indistinguishability, if and only if $f(b') \models_{\boldsymbol{A}} \alpha$, again by the infomorphism condition.

4.3. The condition is that for every type $\alpha \in \mathrm{typ}(\boldsymbol{A})$ there is a type $\beta \in \mathrm{typ}(\boldsymbol{B})$ such that tok(α) $= g\,[\mathrm{tok}(\beta)]$.

4.5. For any token $b \in \mathrm{tok}(\boldsymbol{B})$, $f(b)$ and $g(b)$ have the same types, by the infomorphism condition, and so $f(b) = g(b)$, because $\boldsymbol{A}$ is separated.

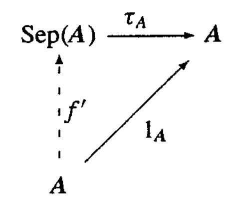
4.7. Let $\mathbf{1}$ be a classification with a single type and no tokens. Then for every $\boldsymbol{A}$ there is a unique infomorphism $f : \boldsymbol{A} \rightleftarrows \mathbf{1}$.

Lecture 5

5.1. The isomorphism between $\boldsymbol{A} + \mathbf{o}$ and $\boldsymbol{A}$ is the identity on types. On tokens it takes $\langle a, u\rangle$ to a, where u is the unique token of $\boldsymbol{A}$. The

isomorphism between $\boldsymbol{A} + \boldsymbol{B}$ and $\boldsymbol{B} + \boldsymbol{A}$ is defined on types by taking $\langle 0, \alpha\rangle$ to $\langle 1, \alpha\rangle$ and $\langle 1, \beta\rangle$ to $\langle 0, \beta\rangle$. On tokens it takes $\langle a, b\rangle$ to $\langle b, a\rangle$. It is easy to see that this is an isomorphism. Associativity is similar, if notationally messy.

5.2. Note that if $\boldsymbol{A}$ is separated, then the identity infomorphism on $\boldsymbol{A}$ respects the invariant $\langle \mathrm{typ}(\boldsymbol{A}), \sim_{\boldsymbol{A}}\rangle$. Thus, by Proposition 5.18, there is a unique infomorphism $f' : \boldsymbol{A} \rightleftarrows \mathrm{Sep}(\boldsymbol{A})$ making the following diagram commute:



But that implies that f' is the inverse of $\tau_{\boldsymbol{A}}$, and so $\tau_{\boldsymbol{A}}$ is an isomorphism. The converse is obvious.

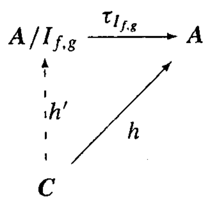
5.3. Suppose that $a_1 \models_{\boldsymbol{A}} \alpha$ and $R_{f,g}(a_1, a_2)$ and $\alpha \in A_{f,g}$. Then there is a token b such that $f(b) = a_1$ and $g(b) = a_2$, and so $b \models_{\boldsymbol{B}} f(\alpha)$ by the infomorphism condition on f. But $f(\alpha) = g(\alpha)$ because $\alpha \in A$, and so $a_2 \models_{\boldsymbol{A}} \alpha$, by the infomorphism condition on g. Hence $I_{f,g}$ is an invariant of $\boldsymbol{A}$.

To show that $h : \boldsymbol{C} \rightleftarrows \boldsymbol{A}$ respects $\langle A_{f,g}, R_{f,g}\rangle$ if and only if $fh = gh$, note the following equivalences:

1. $h(\gamma) \in A_{f,g}$ iff $f^{\wedge}(h^{\wedge}(\gamma)) = g^{\wedge}(h^{\wedge}(\gamma))$
   iff $fh(\gamma) = gh(\gamma)$.
2. $R_{f,g}(a_1, a_2)$ iff $\exists b\; f(b) = a_1$ and $g(b) = a_2$
   iff $\exists b\; fh(b) = h^{\vee}(f^{\vee}(b)) = h^{\vee}(a_1)$ and $gh(b) = h^{\vee}(g^{\vee}(b)) = h^{\vee}(a_2)$.

If $fh = gh$, then by (1) $h(\gamma) \in A_{f,g}$ for all $\gamma \in \mathrm{typ}(\boldsymbol{C})$; and if $R_{f,g}(a_1, a_2)$, then $h(a_1) = h(a_2)$ by (2), as required. If, on the other hand $fh \neq gh$, then either there is a $\gamma \in \mathrm{typ}(\boldsymbol{C})$ such that $f^{\wedge}(h^{\wedge}(\gamma)) \neq g^{\wedge}(h^{\wedge}(\gamma))$ and so by (1) $h(\gamma) \neq A_{f,g}$, or there is a $b \in \mathrm{tok}(\boldsymbol{B})$ such that $fh(b) \neq gh(b)$, so that $h^{\vee}(f^{\vee}(b)) \neq h^{\vee}(g^{\vee}(b))$; but $R_{f,g}(f^{\vee}(b), g^{\vee}(b))$ by definition of $R_{f,g}$, and so h does not respect $\langle A_{f,g}, R_{f,g}\rangle$.

To show that $\tau_{I_{f,g}}$ is the equalizer of f and g, we must show that for every infomorphism $h : \boldsymbol{C} \rightleftarrows \boldsymbol{A}$ for which $fh = gh$ there is a unique infomorphism $h' : \boldsymbol{C} \rightleftarrows \boldsymbol{A}/I_{f,g}$ making the following diagram commute:



Given the equivalence just proved, this is just a restatement of Proposition 5.18.

Finally, for any invariant $I = \langle \Sigma, R \rangle$ of a classification $\boldsymbol{A}$, we can define a classification $\boldsymbol{B}$ and infomorphisms $f_0 : \boldsymbol{A} \rightleftarrows \boldsymbol{B}$ and $f_1 : \boldsymbol{A} \rightleftarrows \boldsymbol{B}$ such that τ_I is the equalizer of f_0 and f_1. First, let $A' = \{\langle i, \alpha \rangle \mid i \in \{0, 1\}, \alpha \in \text{typ}(\boldsymbol{A}) - A\}$ be the disjoint union of $\text{typ}(\boldsymbol{A}) - A$ with itself. Now define $\boldsymbol{B}$ to be the classification whose types are members of $A \cup A'$, whose tokens are the pairs $\langle a_0, a_1 \rangle$ for which $R(a_0, a_1)$, and such that

1. for $\alpha \in A$, $\langle a_0, a_1 \rangle \models_{A'} \alpha$ if and only if $a_0 \models_{A} \alpha$, and
2. for $\langle i, \alpha \rangle \in A'$, $\langle a_0, a_1 \rangle \models_{A'} \langle i, \alpha \rangle$ if and only if $a_i \models_{A} \alpha$.

Define $f_i : \boldsymbol{A} \rightleftarrows \boldsymbol{B}$ by $f_i \langle a_0, a_1 \rangle = a_i$ and

$$f_i(\alpha) = \begin{cases} \alpha & \text{if } \alpha \in A \\ \langle i, \alpha \rangle & \text{if } \alpha \notin A \end{cases}.$$

It is easy to see that $A_{f_0, f_1} = A$ and $R_{f_0, f_1} = R$.

Lecture 6

6.3. Hint: In the first section we showed how to take a sequential composition of channels. The limit construction can also be used to take a parallel composition of channels. Suppose we have two separate channels connecting $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, say $\mathcal{C}_1 = \{f_i : \boldsymbol{A}_i \rightleftarrows \boldsymbol{B_1}\}_{i \in \{1,2,\}}$ and $\mathcal{C}_2 = \{g_i : \boldsymbol{A}_i \rightleftarrows \boldsymbol{B_2}\}_{i \in \{1,2,\}}$. This can be seen as a distributed system with four classifications and four infomorphisms. Determine its minimal cover. Simplify it by obtaining an isomorphic structure that has less redundant structure. Interpret this cover in informational terms.

Lecture 7

7.1. Suppose $\boldsymbol{A}$ is Boolean and that X is closed under indistinguishability. Show that X is represented by the type

$$\alpha = \bigvee_{a \in X} \bigwedge \text{typ}(a).$$

Clearly, every $a \in X$ is of type α. For the converse, suppose that $a' \models_A \alpha$. Then there is some $a \in X$ such that $a' \models_A \bigwedge \text{typ}(a)$. We show that a' is indistinguishable from a and so is in X, because X is closed under indistinguishability. Clearly, every type of a is a type of a'. Suppose $a' \models_A \alpha'$. Then $a' \not\models_A \neg\alpha'$, so $a \not\models_A \neg\alpha'$. Hence $a \models_A \alpha$. Thus a is indistinguishable from a'.

For the converse, suppose that for every set X closed under indistinguishablity there is a type α such that $X = \text{typ}(\alpha)$. We show that $\boldsymbol{A}$ has a conjunction, disjunction, and negation. Let us start by showing negation. Let $\beta \in \text{typ}(\boldsymbol{A})$ be arbitrary. The set $\text{tok}(\boldsymbol{A}) - \text{tok}(\beta)$ is closed under indistinguishablity so it is definable by some type. Choose one and let that be the negation of β. This clearly has the desired properties. Now suppose that Θ is an arbitrary set of types. Let $X = \bigcup_{\beta \in \Theta} \text{tok}(\beta)$. Clearly, X is closed under indistinguishablity. Hence it is definable by some type α. Pick such an α and let that be the disjunction of Θ. Conjunction is similar or can be defined in terms of disjunction and negation.

Lecture 8

8.1. Let $\boldsymbol{A} = \boldsymbol{B} = \text{Evt}(\boldsymbol{S}_{DNA})$, and let $\boldsymbol{C}$ be the restriction of $\text{Evt}(\boldsymbol{S}_{DNA})$ to those tokens that have successfully split. Let $f : \boldsymbol{A} \rightleftarrows \boldsymbol{C}$ and $b : \boldsymbol{B} \rightleftarrows \boldsymbol{C}$ be type-identical infomorphisms that on tokens are defined so that $f(c)$ is one of c's offspring and $g(c)$ is the other. The infomorphism condition on c is the condition that the splitting is normal, that is, that no mutation takes place.

8.2. If $f : \boldsymbol{S}_1 \rightleftarrows \boldsymbol{S}_2$ is an isomorphism, then $\langle f^{\wedge}, f^{\vee -1} \rangle$ is a projection satisfying the above conditions. Conversely, if $f : \boldsymbol{S}_1 \rightrightarrows \boldsymbol{S}_2$ is a projection and both $f^{\wedge}$ and $f^{\vee}$ are bijections, we show that $\langle f^{\wedge}, f^{\vee -1} \rangle$ is an isomorphism. The only thing that needs to be checked is that it is an infomorphism. Let $b \in \text{tok}(\boldsymbol{S}_2)$ and $\alpha \in \text{typ}(\boldsymbol{S}_1)$. The following are equivalent:

$$\begin{array}{llll} f(b) \models_{S_1} \alpha & \text{iff} & \text{state}_{S_1}(f(b)) = \alpha & \text{(by definition of } \mathit{states}_{S_1}) \\ & \text{iff} & f(\text{state}_{S_2}(b)) = \alpha & \text{(because } f \text{ is a projection)} \\ & \text{iff} & \text{state}_{S_2}(b) = f^{-1}(\alpha) & \text{(because } f^\wedge \text{ is a bijection)} \\ & \text{iff} & b \models_{S_2} f^{-1}(\alpha) & \text{(by definition of } \mathit{states}_{S_2}) \end{array}$$

8.4. Here f is the identity on types and is $f(s) = \text{state}_S(s)$ on tokens s.

Lecture 9

9.1. The direction from left to right is immediate; if T is inconsistent, then $\Gamma \vdash_T \Delta$ for all sequents. The other direction follows from the fact that regular theories satisfy Weakening.

9.2. The reflexivity of $\leq_T$ follows from Identity. The transitivity follows from Finite Cut. Saying that T is algebraic is equivalent to

if $\alpha \vdash_T \beta$ and $\beta \vdash_T \alpha$ then $\alpha = \beta$.

If $T = \text{Th}(A)$, then this is equivalent to

$\text{tok}(\alpha) \subseteq \text{tok}(\beta)$ and $\text{tok}(\beta) \subseteq \text{tok}(\alpha)$ implies $\alpha = \beta$.

But this is equivalent to

$\text{tok}(\alpha) = \text{tok}(\beta)$ implies $\alpha = \beta$,

which is the definition of extensionality.

9.3. Define $\Gamma \vdash_T \Delta$ if and only if $\alpha \leq \beta$ for some $\alpha \in \Gamma$ and some $\beta \in \Delta$. Verify that this theory is regular.

9.4. (1) The consequence relation is given by

$$\Gamma \vdash \Delta \quad \text{iff} \quad \Gamma \cap \Delta \neq \emptyset$$

for all sequents $\langle \Gamma, \Delta \rangle$ of Σ. Check that this is a regular theory. It is clearly contained in every such theory. Its ordering $\leq_T$ is the identity relation on Σ, which is a partial ordering, so the theory is algebraic. (2) The largest regular consequence relation is simply the set of all sequents. It is not algebraic if there is more than one type in Σ.

9.5. Let $\text{state}_A(a) = \langle \Gamma_a, \Delta_a \rangle$. The following are equivalent:

a is a counterexample to $\langle \Gamma, \Delta \rangle$	
$\Gamma \subseteq \Gamma_a$ but $\Delta \cap \Gamma_a = \emptyset$	(defn. satisfaction)
$\Gamma \subseteq \Gamma_a$ and $\Delta \subseteq \Delta_a$	($\langle \Gamma_a, \Delta_a \rangle$ a partition)
$\langle \Gamma, \Delta \leq \langle \Gamma_a, \Delta_a \rangle$	(defn. $\leq$)

9.8. The direction from left to right is immediate. For the converse, suppose $\vdash$ is closed under Identity, Weakening, and Finite Cut. Show that the relation satisfies Partition and hence Global Cut. To show Partition, show that every consistent sequent can be extended to a consistent partition. Suppose that $\langle \Gamma, \Delta \rangle$ is consistent. Let X be the set of consistent extensions of $\langle \Gamma, \Delta \rangle$. By compactness, $\langle X, \leq \rangle$ is closed under unions of chains, so, by Zorn's Lemma, it has a maximal element, say $\langle \Gamma', \Delta' \rangle$. This sequent must be a consistent partition extending $\langle \Gamma, \Delta \rangle$. To see this, note that $\Gamma' \cap \Delta' = \emptyset$ by Identity. To see that $\Gamma' \cup \Delta' = \Sigma$, suppose $\alpha \in \Sigma$ but $\alpha \notin \Gamma' \cup \Delta'$. By maximality, the sequents $\langle \Gamma', \Delta' \cup \{\alpha\} \rangle$ and $\langle \{\alpha\} \cup \Gamma', \Delta' \rangle$ are constraints. In other words, $\Gamma' \vdash \Delta', \alpha$ and $\Gamma', \alpha \vdash \Delta'$. Then by Finite Cut, $\Gamma' \vdash \Delta'$, contradicting the consistency of $\langle \Gamma', \Delta' \rangle$.

9.9. It is straightforward to check that $\vdash$ satisfies Identity, Weakening, and Finite Cut. To prove that it is not regular, let $\Gamma = \Delta = \emptyset$. Notice that every partition of Σ extending $\langle \Gamma, \Delta \rangle$ is a constraint of $\vdash$. By Partition, $\Gamma \vdash \Delta$. But this violates the definition of $\vdash$.

9.10. Part (1) follows easily using Partition. For Part (2), suppose $\Gamma, \Gamma' \vdash \Delta$ and for each $\alpha \in \Gamma'$, $\Gamma \vdash \Delta, \alpha$ but $\Gamma \nvdash \Delta$. Then $\Gamma \cap \Delta = \emptyset$ and $\Gamma \cup \Delta$ has an infinite complement. Thus either Γ' contains infinite many types not in $\Gamma \cup \Delta$ or it contains an element $\alpha_0 \in \Delta$. The latter can't be true because $\Gamma \vdash \Delta, \alpha_0$ and the right side of this is just Δ. But if the former is the case, pick some such $\alpha_1 \in \Gamma'$. Then $\Gamma \vdash \Delta, \alpha_1$. But this cannot be because both sides are disjoint but the complement of their union is still infinite. Part (3) follows immediately from parts (1) and (2) and the fact that this consequence relation is not regular.

9.12. This is an easy consequence of the observation that $\alpha \not\leq_T \beta$ if and only if there is a consistent partition $\langle \Gamma, \Delta \rangle$ with $\alpha \in \Gamma$ and $\beta \in \Delta$.

9.15. Define $f^{*\wedge} = f$ and $f^{*\vee}(a) = \langle f^{-1}[\Gamma_a], f^{-1}[\Delta_a] \rangle$ where $\langle \Gamma_a, \Delta_a \rangle$ is state description of a in A.

Lecture 10

10.1. (1) If $A \sqsubseteq B$, then $\text{typ}(A) \subseteq \text{typ}(B)$. Also, if $\langle \Gamma, \Delta \rangle$ is a sequent of $\text{typ}(A)$ and $\Gamma \nvdash_B \Delta$, then $\langle \Gamma, \Delta \rangle$ has a counterexample $b \in \text{tok}(B)$, so b must also be a counterexample to $\langle \Gamma, \Delta \rangle$ in A because $A \sqsubseteq B$; hence $\Gamma \nvdash_A \Delta$. (2) If the converse of (1) holds, then $A = B$ if and only if $\text{Th}(A) = \text{Th}(B)$. But that can't be. Just take some nontrival classification A and a similar classification with an additional token that is indistinguishable from some token of A. They have the same theories.

Lecture 11

11.1. Let f be type identical and on tokens be defined by

$$f(\langle \Gamma, \Delta) = \langle \Delta, \Gamma \rangle.$$

It is easy to check that this is an isomorphism.

11.4. These results are easy consequences of the rules for the connectives. For negation, what follows is that $\neg\neg\alpha = \alpha$ and that $\alpha \leq_T \beta$ if and only if $\neg\beta \leq_T \neg\alpha$.

Lecture 12

12.1. (1) It is clear that every normal token of $\mathfrak{L} \,|\, \Theta$ satisfies every constraint of $\mathfrak{L} \,|\, \Theta$. We also need to check that the theory of $\mathfrak{L} \,|\, \Theta$ is regular. Weakening is clear. To show Partition, suppose $\Gamma \nvdash_{\mathfrak{L}|\Theta} \Delta$. Then $\Gamma, \Theta \nvdash_{\mathfrak{L}} \Delta$. Let $\langle \Gamma', \Delta' \rangle$ be a consistent partition of $\text{typ}(\mathfrak{L})$ extending $\langle \Gamma \cup \Theta, \Delta \rangle$. Then $\langle \Gamma' - \Theta, \Delta' \rangle$ is an extension of $\langle \Gamma, \Delta \rangle$ that is a consistent partition in $\mathfrak{L} \,|\, \Theta$. (3) Let θ be a new type not in $\text{typ}(\mathfrak{L})$ and let A be the classification that is just like $\text{cla}(\mathfrak{L})$ except that A has the additional type θ with $\text{typ}(\theta) = N_{\mathfrak{L}}$. Our sound local logic will have A for its classification. Because the logic is to be sound, we need only specify its theory. Let $\Gamma \vdash_{\mathfrak{L}'} \Delta$ if and only if one of the following conditions hold:

1. $\theta \notin \Gamma$ and $\Gamma \cap \Delta \neq \emptyset$.
2. $\theta \in \Gamma \cap \Delta$.
3. $\theta \in \Gamma - \Delta$ and $\Gamma - \{\theta\} \vdash_{\mathfrak{L}} \Delta$.

Clearly every token of A satisfies all these constraints; they satisfy the constraints in (3) because $\mathfrak{L}$ is a local logic and $\text{typ}(\theta) = N_{\mathfrak{L}}$. So

we need only verify that this is indeed a regular theory. We check Partition, leaving Weakening to the reader. Suppose $\Gamma \nvdash_{\mathfrak{L}'} \Delta$. There are two cases to consider. If $\theta \notin \Gamma$, then $\Gamma \cap \Delta = \emptyset$ by (1). Let $\Gamma' = \Gamma$ and $\Delta' = \text{typ}(\mathfrak{L}') - \Gamma$. Then $\langle \Gamma', \Delta' \rangle$ is a consistent partition extending $\langle \Gamma, \Delta \rangle$. So suppose $\theta \in \Gamma$. Then $\theta \notin \Delta$ by (2) and $\Gamma - \{\theta\} \nvdash_{\mathfrak{L}} \Delta$ by (3). Because $\text{th}(\mathfrak{L})$ is regular, there is a consistent partition $\langle \Gamma', \Delta' \rangle$ of $\mathfrak{L}$ extending $\langle \Gamma - \{\theta\}, \Delta \rangle$. But then $\langle \Gamma' \cup \{\theta\}, \Delta' \rangle$ is a consistent partition of $\mathfrak{L}'$ extending $\langle \Gamma, \Delta \rangle$. Hence every consistent sequent of $\mathfrak{L}'$ has a consistent partition extending it. It is easy to check that $\mathfrak{L} = \mathfrak{L} \,|\, \theta$.

12.2. Let us write $\Gamma \vdash \Delta$ if and only if one of the conditions (a)–(c) holds. We first check that this is a regular consequence relation. It clearly satisfies Weakening. To see that it satisfies Partition, suppose that $\Gamma \nvdash \Delta$. We want to find a consistent partition $\langle \Gamma', \Delta' \rangle$ extending $\langle \Gamma, \Delta \rangle$, that is, such that $\Gamma' \nvdash \Delta'$. By (a), Γ and Δ are disjoint. By (b), Γ contains at most one state. There are now two cases to consider, depending on whether Γ is empty or a singleton. If $\Gamma = \{\sigma\}$, then we let $\Gamma' = \Gamma$ and let $\Delta' = \Omega - \{\sigma\}$, and we have our desired consistent partition. If $\Gamma = \emptyset$, then we apply (c) to find some $\sigma \notin \Delta$ and let $\Gamma' = \{\sigma\}$ and let $\Delta' = \Omega - \{\sigma\}$, again finding a consistent partition. This shows that $\vdash$ is a regular consequence relation. It is clear that every token satisfies every constraint of this relation, so we have a sound logic. Hence the logic it must contain the logic $\text{Triv}(S)$. On the other hand, by Weakening, each of the sequents above is a constraint of $\text{Triv}(S)$.

12.4 The logic $\mathtt{1}$ is $\text{Log}(\mathtt{1})$, where $1 = \mathbf{o}^{\perp}$.

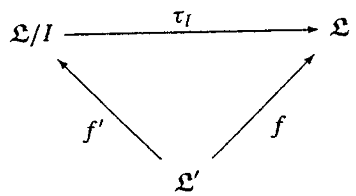
12.3 In each case the bijection between Hom-sets is obvious. For the left adjointness, one just needs to check that an infomorphism $f : A \rightleftarrows \text{cla}(\mathfrak{L})$ preserves constraints and so becomes a logic infomorphism from $\text{Log}(A)$ to $\mathfrak{L}$; this uses the completeness of $\mathfrak{L}$. For the right adjointness, we need to know that the pair of identity functions on $\text{tok}(\mathfrak{L})$ and $\text{typ}(\mathfrak{L})$ preserves normal tokens and so is a logic infomorphism from $\mathfrak{L}$ to $\text{Log}(\text{cla}(\mathfrak{L}))$; this uses the soundness of $\mathfrak{L}$.

12.7. The direction from left to right is immediate. That from right to left is almost as immediate. It is basically the proof of Proposition 10.3.

12.9. First prove that if $B_0 \sqsubseteq B_1$, then $\text{typ}(B_1) \subseteq \text{typ}(B_0)$. Use this, plus Exercise 1, to prove the result.

12.10. Let $\mathfrak{L}$ be a local logic on a classification A of types and let J be an invariant on A.

1. The logic $\mathfrak{L}/I$ is the largest logic on A/I such that the function τ_I is a logic infomorphism.
2. Let $f : \mathfrak{L} \rightleftarrows \mathfrak{L}'$ be a logic infomorphism that respects I. There is a unique logic infomorphism $f' : \mathfrak{L}' \rightleftarrows \mathfrak{L}/I$ such that the following diagram commutes:



12.12. This is just definition unwinding.

Lecture 13

13.1. Let f be the unque infomorphism from $\mathbf{o}$ to A. Then $\mathrm{AP}(A) = f[\mathbf{o}]$, where this $\mathbf{o}$ is the zero logic.

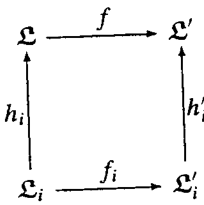
Lecture 15

15.1. Sketch: The information system has the same shape as that used to illustrate the Xerox principle in Lecture 6. We need five classifications, three for classifying molecules and two for classifying processes. First, we need the classification $\mathrm{Evt}(S_{DNA})$ introduced in Example 8.16 We need a second classification for RNA molecules and a third for classifying protein molecules. As for the process classifications, we need one to classify events where a piece of DNA gives rise to a molecule of messenger RNA (notice the informational vocabulary) and one to classify the process whereby an RNA molecule gives rise to a protein. We can turn this into an information system by imposing a logic on each of these classifications, namely, the logics that represent our current understanding of these classifications. Taking the distributed logic of the system gives us a theory of information flow linking all five. As for the one-way flow posulated by the fundamental dogma, we can see that complete information about the strand of DNA gives rise to complete information about the molecule of RNA but that the converse is not the case and similarly, complete information about the molecule of RNA gives rise to complete information about the protein, but the converse is not the case.

15.2. **Definition A.1.** A *logic channel* $\mathcal{C}$ is an indexed family $\{h_i^c : \mathfrak{L}_i \rightleftarrows \mathrm{core}(\mathcal{C})\}_{i \in I}$ of logic infomorphisms with a common codomain, called the *core* of $\mathcal{C}$.

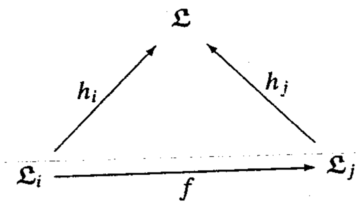
Definition A.2.

1. Given an index set I and logic channels $\mathcal{C} = \{h_i : \mathfrak{L}_i \rightleftarrows \mathfrak{L}\}_{i \in I}$ and $\mathcal{C}' = \{h_i' : \mathfrak{L}_i' \rightleftarrows \mathfrak{L}'\}_{i \in I}$, a *logic channel infomorphism* $f : \mathcal{C} \rightleftarrows \mathcal{C}'$ consists of logic infomorphisms $f : \mathfrak{L} \rightleftarrows \mathfrak{L}'$ and $f_i : \mathfrak{L}_i \rightleftarrows \mathfrak{L}_i'$ such that for each $i \in I$, the following diagram commutes:



If $\mathfrak{L}_i = \mathfrak{L}_i'$ and $f_i = 1_{\mathfrak{L}_i}$ for each $i \in I$, then f is called a *logic refinement infomorphism* and $\mathcal{C}$ is said to be a *logic refinement* of $\mathcal{C}'$.

2. A logic channel $\mathcal{C} = \{h_i : \mathfrak{L}_i \rightleftarrows \mathfrak{L}\}_{i \in I}$ *covers* an information sys tem $\mathcal{L}$ if $\log(\mathcal{L}) = \{\mathfrak{L}_i\}_{i \in I}$ and for each $i, j \in I$ and each infomorphism $f : \mathfrak{L}_i \rightleftarrows \mathfrak{L}_j$ in $\inf(\mathcal{L})$, the following diagram com mutes:



$\mathcal{C}$ is a *minimal cover* of $\mathcal{L}$ if it covers $\mathcal{L}$ and for every other channel $\mathcal{D}$ covering $\mathcal{L}$ there is a unique logic refinement infomorphism from $\mathcal{C}$ to $\mathcal{D}$.

Theorem A.3. *Every information system has a minimal cover. It is unique up to isomorphism.*

To prove this, let $\mathcal{L}$ be an information system with the indexed family $\log(\mathcal{L}) = \{\mathfrak{L}_i\}_{i\in I}$ of local logics together with the set $\inf(\mathcal{L})$ of logic infomorphisms. From it, extract the obvious distributed system with classifications $\mathrm{cla}(\mathfrak{L})_i$ and infomorphisms $\mathrm{Cla}(f)$, for $f \in \inf(\mathcal{L})$. Let $\mathcal{C} = \{h_i^c : \mathrm{cla}(\mathfrak{L})_i \rightleftarrows C\}_{i\in I}$ be the information channel that is the minimal cover of this system. We turn this into a logic channel by taking $\mathfrak{L} = \bigsqcup_{i\in I} h_i^c[\mathfrak{L}_i]_i$. By earlier results, this is the smallest logic on C making all of the infomorphisms h_i^c into logic infomorphisms. It is easy to verify that this is the minimal cover of the information system. Uniqueness is routine.

Lecture 16

16.2. This proof is just definition unwinding.

16.3. We prove (1) and leave (2) to the reader. Both state spaces are subspaces of $\mathrm{Ssp}(A)$, so we need only check that they have the same sets of tokens and same states. It is clear that both state spaces have as sets of tokens the set $f^{\vee}[N_{\mathfrak{L}_2}]$, so we need only check that they have the same set of states. The states of $\mathrm{Ssp}(f^{-1}[\mathfrak{L}_2])$ consist of those partitions of $\mathrm{typ}(A)$ that are consistent in $f^{-1}[\mathfrak{L}_2]$. These are exactly the set of partitions of the form $\langle f^{-1}[\Gamma_2], f^{-1}[\Delta_2]\rangle$ for some consistent partition $\langle\Gamma_2, \Delta_2\rangle$ of $\mathfrak{L}_2$. These, in turn, are exactly the states of $\mathrm{Ssp}(f)[\mathrm{Ssp}(\mathfrak{L}_2)]$.

Lecture 18

18.1. The answer is "No" if the building and person are both tokens of the the same height classification and the building is taller than the person. However, if we allow that the classification of buildings by height is a different classification than that of people by height, then there is no problem.

18.2. Let $r = \langle .1, [0, 48), [48, 60), [60, +\infty)\rangle$. This is the most natural regimentation for A, given the way the problem was stated.

18.3. The classification is determinate, because $.1 \le .25$, and it is precise because these three intervals exhaust the positive real numbers.

18.4. The sorites number of $\mathrm{Log}_r^{\circ}(A)$ is $(.1)^{-1} \times (60 - 48) + 1 = 121$, so this provides a natural upper bound for the Sorites number of A.

18.5. The sorites number of A depends only on the resulting classification, not on any particular regimentation. Another regimentation that would yield the same classification, the one with the largest possible tolerance, is $r' = \langle .25, [0, 48), [48, 60), [60, +\infty)\rangle$. The sorites number of $\mathrm{Log}_{r'}^{\circ}(A)$ is $(.25)^{-1} \times (60 - 48) + 1 = 49$. To find the Sorites number of A we must minimize the distance between the upper endpoint of I_s and the lower endpoint of I_t, while making the tolerance as large as possible. Let h_1 and h_2 be the heights of the shortest and tallest students of medium height. Thus $h_1 \ge 48$ and $h_2 < 60$. This results in a third regimentation $r'' = \langle .25, [0, h_1), [h_1, h_2], (h_2, +\infty)\rangle$. For this logic the number is the least integer $N > (.25)^{-1} \times (h_2 - h_1)$. Thus, given the information about the shortest and tallest girls of medium height, the sorites number of the classification is $4 \times (58.5 - 49) + 1 = 39$.

Lecture 21

21.1. The idealized tokens of the fire fly manual and the penny manual correspond to the rows of the following tables, respectively:

FRONT	SIDE
L	F
L	B
R	F
R	B
N	N

FLIP	BANG
H	H
T	T

21.2. This is a consequence of Theorem 21.10 Let c be an idealized token of type $\Omega_E - \sigma$ and not of type $\Omega_F - \sigma$. In other words, $c_F \in \sigma$ but $c_E \notin \sigma$. However, this contradicts the fact that c is an idealized token, because σ is a subset of both Ω_E and Ω_F.

Bibliography

Auster, P. (1982). *The Invention of Solitude*. London and Boston: Faber and Faber.

Austin, J. L. (1961). How to talk: Some simple ways. In *Philosophical Papers*. J. O. Urmson and G. J. Warnock, eds. Oxford University Press, Oxford.

Barr, M. (1979). **-Autonomous Categories*. Lecture Notes in Mathematics, 752. Springer-Verlag, Heidelberg.

Barr, M. (1991). *-Autonomous categories and linear logic. *Mathematical Structures in Computer Science*, **1**, 159–178.

Barwise, J. (1983). Information and semantics. *The Behavioral and Brain Sciences*, **6**, 65.

Barwise, J. (1986). Information and circumstance: A reply to Fodor. *Notre Dame Journal of Formal Logic*, **27**(3), 324–338.

Barwise, J. (1989). *The Situation in Logic*. Number 17, CSLI Lecture Notes. CSLI Publications, Stanford, Calif.

Barwise, J. (1991). Information links in domain theory. In *Proceedings of the Mathematical Foundations of Programming Semantics Conference (1991)*, S. Brookes, ed., pp. 168–192. Springer-Verlag, Heidelberg.

Barwise, J. (1993). Constraints, channels, and the flow of information. In *Situation Theory and Its Applications, II*, Number 22, CSLI Lecture Notes. P. Aczel et al., eds., pp. 3–27. CSLI Publications, Stanford, Calif.

Barwise, J., and J. Etchemendy (1990). Information, infons, and inference. In *Situation Theory and Its Applications, I*, Number 22, CSLI Lecture Notes. R. Cooper, K. Mukai, and J. Perry, eds., pp. 33–78. CSLI Publications, Stanford, Calif.

Barwise, J., and J. Etchemendy (1995). *Hyperproof*. Number 42, CSLI Lecture Notes. CSLI Publications, Stanford, Calif.

Barwise, J., and J. Perry (1983). *Situations and Attitudes*. MIT Press, Cambridge, Mass.

Barwise, J., and J. Seligman (1994). The rights and wrongs of natural regularity. In *Perspectives in Philosophy, 8*, J. Tomberlin, ed., pp. 331–364. Ridgeview, Atascadero, Calif.

Barwise, J., D. Gabbay, and C. Hartonas (1996). Information flow and the Lambek calculus. In Logic, Language and Computation. Number 26, CSLI Lecture Notes. J. Seligman and D. Westerståhl, eds., pp. 47–62. CSLI Publications, Stanford, Calif.

Birkhoff, G. (1940). *Lattice Theory*. Colloquium Publications. American Math Society, Providence, R.I.

Casti, J. (1992). *Reality Rules, I*. Wiley Interscience, New York.

Cohen, D. W. (1989). *An Introduction to Hilbert Space and Quantum Logic*. Problem Books in Mathematics. Springer-Verlag, New York.

Davey, B., and H. Priestley (1990). *Introduction to Lattices and Order*. Cambridge Mathematical Textbooks, Cambridge, U.K.

Devlin, K. (1991). *Logic and Information*. Cambridge, U.K.: Cambridge University Press.

Dretske, F. (1970). Epistemic operators. *Journal of Philosophy*, **67**, 1007–1023.

Dretske, F. (1971). Conclusive reason. *Australasian Journal of Philosophy*, **49**, 1–22.

Dretske, F. (1981). *Knowledge and the Flow of Information*. MIT Press, Cambridge, Mass.

Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, **43**, 181–4.

Goldman, A. (1979). What is justified belief? In *Justification and Knowledge*. G. Papas, ed. 1–23. Reidel, Dordrecht.

Goldman, A. (1986). *Epistemology and Cognition*. Harvard University Press, Cambridge, Mass.

Hardegree, G. (1982). An approach to the logic of natural kinds. *Pacific Phil. Quarterly*, **63**, 122–132.

Hartonis, C., and J. M. Dunn (1993). Duality theorems for partial orders, semilattics, galois connections and lattices. Technical report, #IULG-93-26. Preprint Indiana University Logic Group. Bloomington, IN.

Hayes, P. (1985). Naive physics, i, Ontology for liquids. In *Formal Theories of the Commonsense World*, R. Hobbs and R. Moore, eds., pp. 71–108. Ablex Press, Norwood, New Jersey.

Kochen, S., and E. Specker (1967). The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, **17**, 59–87.

Koons, R. (1996). Information, representation, and the possibility of error. In *Logic, Language and Computation*. Number 58, CSLI Lecture Notes, J. Seligman and D. Westerståhl, eds., pp. 333–346. CSLI Publications, Stanford, Calif.

Lewis, D. (1973). *Counterfactuals*. Harvard University Press, Cambridge, Mass.

Nozick, R. (1981). *Philosophical Explanations*. Harvard University Press, Cambridge, Mass.

Rabin, M. (1974). Decidable theories. In *Handbook of Mathematical Logic*. J. Barwise, ed. North Holland, Amsterdam.

Seligman, J. (1990). Perspectives in situation theory. In *Situation Theory and Its Applications, I*, Number 22, CSLI Lecture Notes. R. Cooper, K. Mukai, and J. Perry, eds., pp. 147–191. CSLI Publications. Stanford, Calif.

Seligman, J. (1991a). *Perspectives: A Relativistic Approach to the Theory of Information*. Ph.D. thesis, University of Edinburgh.

Seligman, J. (1991b). Physical situations and information flow. In *Situation Theory and Its Applications, II*, Number 26, CSLI Lecture Notes, J. Barwise et al., eds. pp. 257–292. CSLI Publications, Stanford, Calif.

Shannon, C. (1948). The mathematical theory of communication. *Bell System Technical Journal, July and October*, Volume 27, pp. 37–423; 623–656.

Shimojima, A. (1996). *On the Efficacy of Representation*. Ph.D. thesis, Indiana University, Bloomington, IN.

Shope, R. (1983). *The Analysis of Knowing*. Princeton University Press, Princeton, N.J.

Stalnaker, R. (1984). *Inquiry*. MIT Press, Cambridge, Mass.

Swain, M. (1981). *Reasons and Knowledge*. Cornell University Press, Ithaca, N.Y.

Glossary of Notation

AP(A): The *a priori* logic on the classification A.
Boole(A): The Boolean closure of the classification A.
Cha($\mathfrak{L}$): The binary channel representing the local logic $\mathfrak{L}$, a channel relating the classification cla($\mathfrak{L}$) with the idealization Idl($\mathfrak{L}$) of $\mathfrak{L}$.
cla($\mathcal{A}$): The classifications of the distributed system $\mathcal{A}$.
cla($\mathfrak{L}$): The classification of the local logic $\mathfrak{L}$.
Cla(T): The classification generated by the regular theory T.
Cla(f): The infomorphism associated with the theory interpretation f.
Cmp($\mathfrak{L}$): The completion of the local logic $\mathfrak{L}$.
core(C): The core classification of a channel C.
DLog$_C$($\mathfrak{L}$): The logic obtained by distributing the logic $\mathfrak{L}$ on the core of the channel C to its component classifications.
DLog(S): The logic obtained by distributing the logic on the core of the state space S to its projected state spaces.
Evt(S): The event classifications generated by the state space S.
Evt(f): The infomorphism on event classifications corresponding to a projection f of the generating state spaces.
Idl($\mathfrak{L}$): The idealization classification of the local logic $\mathfrak{L}$.
inf($\mathcal{A}$): The infomorphisms of the distributed system $\mathcal{A}$.
lim $\mathcal{A}$: The limit channel of the distributed system $\mathcal{A}$.
Lind($\mathfrak{L}$): The Lindenbaum logic of $\mathfrak{L}$.
Log($\mathcal{L}$): The systemwide logic of an information system.
Log(A): The local logic generated by the classification A.
Log$_C$ (D): The local logic on the distal classification D generated by the binary channel C.
Log(f): The logic infomorphism naturally associated with an infomorphism, projection, or interpretation f.
Log($\mathcal{L}$): The systemwide logic of an information system $\mathcal{L}$.
Log(S): The local logic generated by the state space S.
Log(T): The local logic generated by the regular theory T.
SC($\mathfrak{L}$): The sound completion the local logic $\mathfrak{L}$.
Sep(A): The separated quotient of A, obtained by identifying indistinguishable tokens.
Snd($\mathfrak{L}$): The sound part of the local logic $\mathfrak{L}$.
Ssp(A): The free state space generated by a classification A.
Ssp(f): The state-space projection generated by the infomorphism f.
Ssp($\mathfrak{L}$): The state space generated by the logic $\mathfrak{L}$.
state$_S$ (a): The state of the object a in the state space S.
state$_A$ (a): The state description of a in the classification A.
Th(A): The theory of the classification A.
Th(f): The interpretation associated with an infomorphism f or a projection.
tok(A): The set of tokens of the classification A.
Triv(S): The trivial logic on a state space.
typ(A): The set of types of the classification A.
$\models_A$: The classification relation of A.

$f: A \rightleftarrows B$: f is a contravariant pair of functions, typically an infomorphism.
$f^\wedge$: The type part of an infomorphism or projection.
$f^\vee$: The token part of an infomorphism or projection.
1_A: The identity infomorphism of the classification A.
h_i^C: The ith infomorphism of the channel C.
$A^\perp$: The flip dual of the classification A, obtained by interchanging types and tokens.
$f^\perp$: The flip dual of the infomorphism f.
$A + B$: The sum of the classifications A and B.
σ_A: The infomorphism $\sigma_A : A \rightleftarrows A + B$ from the classification A into the sum $A + B$.
A/I: The quotient of A by the invariant I. The notation A/J is likewise used for the dual notion when J is a dual invariant.
$A \upharpoonright \Sigma$: The restriction of the classification A to the types in Σ.
τ_I: The canonical quotient infomorphism $\tau_I : A/I \rightleftarrows A$.
$\vee A$: The disjunctive power classification of the classification A.
$\wedge A$: The conjunctive power classification of the classification A.
η_A^d: The natural embedding infomorphism of A into its disjunctive power $\vee A$.
η_A^c: The natural embedding infomorphism of A into its conjunctive power $\wedge A$.
$\neg A$: The negation of the classification A.
$\vee f$: The lifting of $f: A \rightleftarrows B$ to $\vee f: \vee A \rightleftarrows \vee B$.
$\wedge f$: The lifting of $f: A \rightleftarrows B$ to $\wedge f: \wedge A \rightleftarrows \wedge B$.
$\neg f$: The lifting of $f: A \rightleftarrows B$ to $\neg f: \neg A \rightleftarrows \neg B$.
$f: S_1 \rightrightarrows S_2$: A covariant pair of functions f from one state space to another, usually a projection.
$S_1 \times S_2$: The product of state spaces S_1 and S_2.
π_{S_i}: The projection from a product state space onto its ith factor.
$T + T'$: The sum of theories T and T'.
$\sqsubseteq$: The inclusion ordering on theories.
$T \upharpoonright \Sigma$: The restriction of the theory T to the set Σ of types.
T/R: The quotient of the theory T by the relation R.
$f^{-1}[T]$: The inverse image of the theory T under the function f.
$f[T]$: The image of the theory T under the function f.
$\vee T$: The disjunctive power of the theory T.
$\wedge T$: The conjunctive power the theory T.
$\neg T$: The negation of the theory T.
$\vdash_{\mathfrak{L}}$: The consequence relation of the local logic $\mathfrak{L}$.
$N_{\mathfrak{L}}$: The set of normal tokens of the local logic $\mathfrak{L}$.
$\mathfrak{L} \upharpoonright \Theta$: The conditionalization of the local logic $\mathfrak{L}$ on the set Θ of types.
$\mathfrak{L}_1 + \mathfrak{L}_2$: The sum of local logics.
$\mathfrak{L}/I$: The quotient of the local logic $\mathfrak{L}$ by the invariant or dual invariant I
$\mathfrak{L}_1 \sqsubseteq \mathfrak{L}_2$: The information ordering on local logics.
$\mathfrak{L}_1 \sqcup \mathfrak{L}_2$: The join of local logics.
$\mathfrak{L}_1 \sqcap \mathfrak{L}_2$: The meet of local logics.
$f[\mathfrak{L}]$: An image of a local logic.
$f^{-1}[\mathfrak{L}]$: An inverse image of a local logic.
$S_{\mathfrak{L}}$: The subspace of the state space S determined by the $\mathfrak{L}$.

Index of Definitions

Index of Names