Recall that a set E of equations is *inconsistent* if it has only trivial models. The proof of the following corollary is left as an exercise.

Corollary 4.7. *A set E of Σ-equations is inconsistent iff $E \vdash x \approx y$, where x and y are distinct variables.*

This corollary can be used to obtain a new proof of the Compactness Theorem of Equational Logic that does not use reduced products. This is also left as an exercise.

4.1. **Logical consequence as a closure operator.** Recall that formally an equation is defined to be an ordered pair $\langle t, s\rangle$ with $t, s \in \mathrm{Te}_\Sigma(X)$. Thus the set of Σ-equations can be identified with the set $\mathrm{Te}_\Sigma(X)^2$. Define $\mathrm{Cn}_\Sigma : \mathcal{P}(\mathrm{Te}_\Sigma(X)^2) \to \mathcal{P}\big(\mathrm{Te}^2_\Sigma(X)\big)$ by

$$\mathrm{Cn}_\Sigma\big(E\big) = \{\, \varepsilon \in \mathrm{Te}^2_\Sigma(X) : E \vDash \varepsilon \,\}.$$

$\mathrm{Cn}_\Sigma\big(E\big)$ is the set of all logical consequences of E, that is, the set of all equations that are identities of every Σ-algebra in which each equation of E is an identity. By the soundness and completeness theorems, $\mathrm{Cn}_\Sigma(E)$ is the set of all equations that are provable from E:

$$\mathrm{Cn}_\Sigma\big(E\big) = \{\, \varepsilon \in \mathrm{Te}^2_\Sigma(X) : E \vdash \varepsilon \,\}.$$

Recall, that, for any class K of Σ-algebras, $\mathrm{Id}(\mathsf{K})$ is the set of all identities of K, i.e., $\mathrm{Id}(\mathsf{K}) = \{\, \varepsilon \in \mathrm{Te}_\Sigma(X) : \mathsf{K} \vDash \varepsilon \,\}$. Cn_Σ can be expressed in terms of the operators Mod and Id as follows.

$$\mathrm{Cn}_\Sigma(E) = \mathrm{Id}\big(\mathsf{Mod}(E)\big).$$

Cn_Σ is a closure operation on the set of Σ-equations, in fact an algebraic closure relation. This can be proved directly, but it turns out to be a consequence of a general method we now discuss for constructing closure operations in a wide variety of difference situations.

4.1.1. *Galois Connections.*

Definition 4.8. Let $\boldsymbol{A} = \langle A, \leq\rangle$ and $\boldsymbol{B} = \langle B, \leq\rangle$ be posets. Let $h: A \to B$ and $g: B \to A$ be mappings such that for all $a, a' \in A$ and $b, b' \in B$,

(i) $a \leq a'$ implies $h(a) \geq h(a')$.
(ii) $b \leq b'$ implies $g(b) \geq h(b')$.
(iii) $a \leq g\big(h(a)\big)$ and $b \leq g\big(h(b)\big)$.

The mappings h and g are said to define a *Galois connection* between $\boldsymbol{A}$ and $\boldsymbol{B}$.

Example. For sets E and F of Σ-equations and any classes K and L of Σ-algebras we have

- $E \subseteq F$ implies $\mathsf{Mod}(E) \supseteq \mathsf{Mod}(F)$.
- $\mathsf{K} \subseteq \mathsf{L}$ implies $\mathrm{Id}(\mathsf{K}) \supseteq \mathrm{Id}(\mathsf{L})$.
- $\mathsf{K} \subseteq \mathsf{Mod}\big(\mathrm{Id}(\mathsf{K})\big)$ and $E \subseteq \mathrm{Id}\big(\mathsf{Mod}(\mathsf{K})\big)$.

Thus Mod and Id are a Galois connection between the posets, in fact complete lattices, $\big\langle \mathcal{P}\big(Te_\Sigma(X)^2\big), \subseteq\big\rangle$ and $\big\langle \mathcal{P}\big(\mathsf{Alg}(\Sigma)\big), \subseteq\big\rangle$.

Galois connections give rise to closure operators in a natural way; before showing this however we first describe the more general notion of a closure operator on a poset.

Definition 4.9. Let $\boldsymbol{A} = \langle A, \leq\rangle$ be a poset. A map $c: A \to A$ is a *closure operator on* $\boldsymbol{A}$ if, for all $a, a' \in A$,

(i) $a \leq c(a)$;
(ii) $c(c(a)) = c(a)$;
(iii) $a \leq a'$ implies $c(a) \leq c(a')$.

Note that a closure operator C on a set A in the sense of Theorem 1.21 is a closure closure operator on the poset $\langle \mathcal{P}(A), \subseteq \rangle$.

Theorem 4.10. *Let h, g be a Galois connection between the posets $\boldsymbol{A} = \langle A, \leq \rangle$ and $\boldsymbol{B} = \langle B, \leq \rangle$. Then $g \circ h: A \to A$ and $h \circ g: B \to B$ are closure operators on $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively.*

Proof. We verify the three conditions of Defn. 4.9. Let a, a' be arbitrary elements of A and b, b' arbitrary elements of B.

(i) $a \leq (g \circ h)(a)$ and $b \leq (h \circ g)(b)$ by definition of a Galois connection.

(iii) $a \leq a'$ implies $h(a) \geq h(a')$ which in turn implies $(g \circ h)(a) \leq (g \circ h)(a')$. $b \leq b'$ implies $g(b) \geq g(b')$ which in turn implies $(h \circ g)(b) \leq (h \circ g)(b')$.

(ii) By (i), $(g \circ h)(a) \leq (g \circ h)\big((g \circ h)(a)\big)$. Also by (ii), $h(a) \leq (h \circ g)\big(h(a)\big)$. Then $g(h(a)) \geq g\big((h \circ g)(h(a))\big)$, i.e., $(g \circ h)(a) \geq (g \circ h)\big((g \circ h)(a)\big)$. So $(g \circ h)(a) = (g \circ h)(g \circ h)(a)$. Similarly, $(h \circ g)(b) = (h \circ g)(h \circ g)(b)$. □

We now show how every binary relation between two sets induces a Galois connection. Let A and B be sets and $R \subseteq A \times B$. Define $\bar{h}: A \to B$ and $\bar{g}: B \to A$ by

$$\bar{h}(a) = \{\, b \in B : a \, R \, b \,\} \quad \text{and} \quad \bar{g}(b) = \{\, a \in A : \underbrace{a \, R \, b}_{b \, \breve{R} \, a} \,\}.$$

Define $H: \mathcal{P}(A) \to \mathcal{P}(B)$ and $G: \mathcal{P}(B) \to \mathcal{P}(A)$ by

$$H(X) = \bigcap \{\, \bar{h}(x) : x \in X \,\} = \{\, b \in B : \forall x \in X \, (x \, R \, b) \,\} \quad \text{and}$$

$$G(Y) = \bigcap \{\, \bar{g}(y) : y \in Y \,\} = \{\, a \in A : \forall y \in Y \, (a \, R \, y) \,\},$$

See Figure 23.

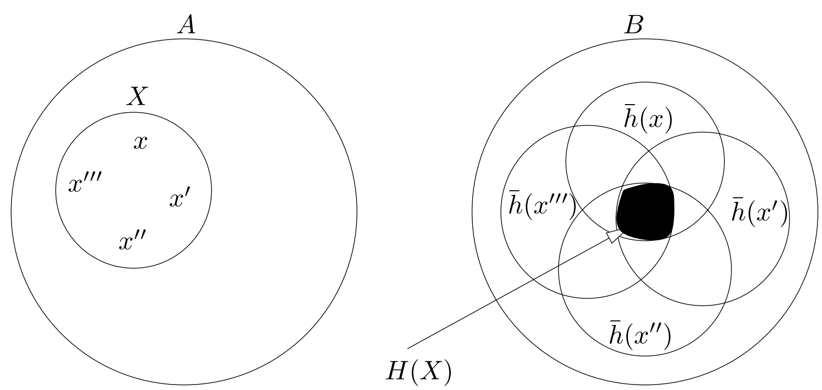

FIGURE 23

Theorem 4.11. *Let A, B be sets and $R \subseteq A \times B$. Then the H and G defined above give a Galois connection between $\langle \mathcal{P}(A), \subseteq \rangle$ and $\langle \mathcal{P}(B), \subseteq \rangle$.*

Proof. Consider any $X, X' \subseteq A$ such that $X \subseteq X'$. Clearly $\{\, \bar{h}(x) : x \in X \,\} \subseteq \{\, \bar{h}(x) : x \in X$, and hence $\bigcap \{\, \bar{h}(x) : x \in X \,\} \supseteq \bigcap \{\, \bar{h}(x) : x \in X$. So $X \subseteq X'$ implies $H(X) \supseteq H(X')$, and similarly, for all $Y, Y' \subseteq B$, $Y \subseteq Y'$ implies $G(Y) \supseteq G(Y')$.

Note that, for all $x \in A$ and $y \in B$,

$$y \in \bar{h}(x) \quad \text{iff} \quad x\,R\,y \quad \text{iff} \quad x \in \bar{g}(y).$$

Thus, for all $x \in A$, $x \in \bigcap\{\, \bar{g}(y) : y \in \bar{h}(x) \,\} = G(\bar{h}(x))$. If $x \in X$, then $H(X) \subseteq \bar{h}(x)$ by definition. So $G(H(X)) \supseteq G(\bar{h}(x))$, and hence for all $x \in X$, $x \in G(H(X))$, i.e., $X \subseteq G(H(X))$. Similarly, for every $Y \subseteq B$, $Y \subseteq H(G(Y))$. □

As a consequence of this theorem and Theorem 4.10, we have that $H \circ G\colon \mathcal{P}(A) \to \mathcal{P}(B)$ and $G \circ H\colon \mathcal{P}(B) \to \mathcal{P}(A)$ are closure operators on A and B respectively in the sense of Theorem 1.21.

Theorem 4.12. *Let A and B be sets and $R \subseteq A \times B$. Let $H\colon \mathcal{P}(A) \to \mathcal{P}(B)$ and $G\colon \mathcal{P}(B) \to \mathcal{P}(A)$ be the Galois connection defined by R. Let $\mathcal{C}_A = \{\, C \subseteq A : (G \circ H)(C) = C \,\}$, the closed subsets of A under $G \circ H$. Let $\mathcal{C}_B = \{\, C \subseteq B : (H \circ G)(C) = C \,\}$, the closed subsets of B under $H \circ G$. The complete lattices $\langle \mathcal{C}_A, \subseteq \rangle$ and $\langle \mathcal{C}_B, \subseteq \rangle$ are dually isomorphic under H.*

The proof is left as an exercise

If we take $A = \mathsf{Alg}(\Sigma)$ (the class of Σ-algebras) and $B = \mathrm{Te}_\Sigma(X)^2$ (the set of Σ-equations), and take

$$R = \vDash = \{\, \langle \boldsymbol{A}, \varepsilon \rangle : \boldsymbol{A} \vDash \varepsilon \,\} \subseteq \mathsf{Alg}(\Sigma) \times \mathrm{Te}_\Sigma(X)^2,$$

Then $H(\mathsf{K}) = \mathrm{Id}(K)$, the set of identities of K, and $G(E) = \mathsf{Mod}(E)$, the variety axiomatized by E. Thus, the consequence operator $\mathrm{Cn}_\Sigma = \mathrm{Id} \circ \mathsf{Mod}$ is a closure operator on the set of Σ-equations. In fact we have

Theorem 4.13. *The Cn_Σ is a finitary closure operation on the set of Σ-equations.*

Proof. As observed above, that Cn_Σ is a closure operator follows from Thms. 4.10 and 4.11. Consider any set E of Σ-equations. By the monotonicity of Cn we have $\bigcup\{\, \mathrm{Cn}(E') : E' \subseteq_\omega E \,\} \subseteq \mathrm{Cn}(E)$. Let $\varepsilon \in \mathrm{Cn}(E)$; $E \vdash \varepsilon$. Let $\delta_1, \dots, \delta_m$ be a proof of ε from E. Obviously there can be only a finite number of applications of the (E-*axiom*) in the proof. Let E' be the finite set of equations in E used in these applications. Then $\delta_1, \dots, \delta_m$ is also a proof of ε from E'. So $E' \vdash \varepsilon$. Hence $E' \vDash \varepsilon$ and $\varepsilon \in \mathrm{Cn}_\Sigma(E')$. So $\mathrm{Cn}_\Sigma(E) \subseteq \bigcup\{\, \mathrm{Cn}(E') : E' \subseteq_\omega E \,\}$. □

The closed-sets, i.e., the sets T of equations such that $\mathrm{Cn}_\Sigma(T) = T$ are called (*equational*) *theories*. T is a theory iff it is closed under consequence, i.e., $T \vDash \varepsilon$ (equivalently $T \vdash \varepsilon$) implies $\varepsilon \in T$. The theories form an algebraic closed-set system, and a complete lattice under set-theoretical inclusion. We will obtain a useful theories below.

Consider the dual closure operator $\mathsf{Mod} \circ \mathrm{Id}$ on $\mathsf{Alg}(\Sigma)$: by the Birkhoff HSP Theorem we have $\mathsf{Mod} \circ \mathrm{Id}(\mathsf{K}) = \mathbf{S}\,\mathbf{H}\,\mathbf{P}(\mathsf{K})$. The closed sets are the varieties. They form an closed-set system and hence are closed intersection (this is easy to verify directly); it is not however algebraic. They form a complete lattice under set-theoretical inclusion that is dually isomorphic to the lattice of theories by Thm. 4.12.

The informal method of proving an equation from a given set of equations can be formalized in a more direct way than we did Defn. 4.3. This is done in the following definition.

Definition 4.14. Let E be a set of Σ-equations. Let $\breve{I} = \{\, t \approx s : (s \approx t) \in E \,\}$. Define a relation $\equiv_E \subseteq \mathrm{Te}_\Sigma(X)^2$ as follows. $t \equiv_E s$ if there exists an equation $u(x_0, \dots, x_{n-1}) \approx$

$v(x_0,\dots,x_{n-1})$ in $E\cup\breve{E}$ and $w_0,\dots,w_{n-1}\in \mathrm{Te}_\Sigma(X)$ such that t has a subterm $u(w_0,\dots,w_{n-1})$ and s is obtained from t be replacing this subterm by $v(w_0,\dots,w_{n-1})$.

Example. Let Σ consist of a single binary operation; as usual, we omit the operation symbol and simply concatenate terms. Suppose E contains the associtive law $(xy)z\approx x(yz)$. Let $t=\Big((xy)\big((uw)z\big)\Big)(xy)$. Let $u(x,y,z)\approx v(x,y,z)$ be $x(yz)\approx(xy)z$ in $\breve{E}$ and let $w_0=xy$, $w_1=xw$, $w_2=z$. Then $u(w_0,w_1,w_2)$ is the subterm $(xy)\big((uw)z\big)$ of t and when this is replaced by $v(w_0,w_1,w_2)=\big((xy)(uw)\big)z$ we get $s=\Big(\big((xy)(uw)\big)z\Big)(xy)$. Thus $\Big((xy)\big((uw)z\big)\Big)(xy)\equiv_E\Big(\big((xy)(uw)\big)z\Big)(xy)$. The process of forming $\equiv_E$ can best be visualized of in terms of manipulating parse trees. See for Figure 24

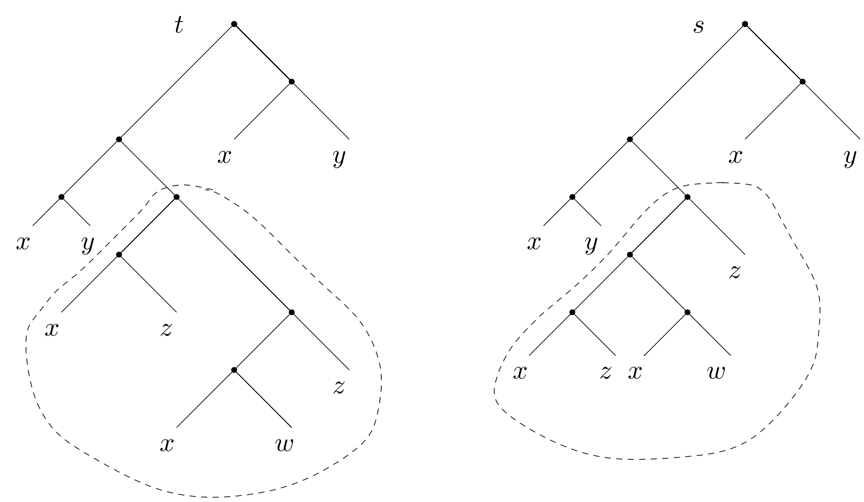

FIGURE 24

Let $\equiv_E^*$ be the the reflexive, transitive closure of $\equiv_E$. $\equiv_E$, i.e., $t\equiv_E s$ if $t=s$ or there exist $r_0,\dots,r_m$ such that $t=r_0\equiv_E r_1\equiv_E\cdots\equiv_E r_m=s$.

Theorem 4.15. *For any set of Σ-equations, $E\vdash t\approx s$ iff $t\equiv_E^* s$.*

Proof. We first prove that, for each E, $\equiv_E^*$ is a congruence relation on the term algebra $\mathbf{Te}_\Sigma(X)$. It is obiously reflexive and transitive. To see it is symmetric, assume $t\equiv_E^* s$. Then s is obtained from t by replacing a subterm of the form $u(\hat{w})$ by $v(\hat{w})$ where $(u(\hat{x})\approx v(\hat{x}))\in E\cup\breve{E}$ and $\hat{w}$ is an arbitrary choice of Σ-terms to substitute for the variables of $\hat{x}$. But clearly $v(\hat{x})\approx u(\hat{x})$ is also in $E\cup\breve{E}$ and t is obtained from s by replacing $v(\hat{w})$ by $u(\hat{w})$. So $s\equiv_E^*$.

We now verify that $\equiv_E^*$ has that substitution property. For this purpose we introduce some useful notation. If $u(\hat{w})$ is a subterm of t, we will denote by $t[v(\hat{w})/u(v(\hat{w})]$ the term that is obtained by replacing $u(\hat{w})$ by $v(\hat{w})$. Suppose $\sigma\in\Sigma_n$ and $t_i\equiv_E^* s_i$ for each $i\le n$. We must show that $\sigma(t_1,\dots,t_n)\equiv_E^*\sigma(s_1,\dots,s_n)$. For each $i\le n$, we have $t_i=r_{i0}\equiv_E r_{i1}\equiv_E\cdots\equiv_E r_{im}=s_i$; we can assume that this sequence is the same length

m for all i by adding repetitions of the last term if needed (this uses that fact that $\equiv_E^*$ is reflexive). By the transitivity of $\equiv_E^*$ it suffices to show for all $i < n$ and all $j < m$ that

$$\sigma(r_{1(j+1)}, \dots, r_{(i-1)(j+1)}, r_{ij}, r_{(i+1)j}, \dots, r_{nj}) \equiv_E \sigma(r_{1(j+1)}, \dots, r_{(i-1)(j+1)}, r_{i(j+1)}, r_{(i+1)j} \dots, r_{nj}).$$

Consequently, without loss of generality we can assume that, for some $i \leq n$, $t_i \equiv_E s_i$ and $t_j = s_j$ for all $j \neq i$. We want to show that

$$\sigma(t_1, \dots, t_{i-1}, t_i, t_{i+1}, \dots, t_n) \equiv_E \sigma(t_1, \dots, t_{i-1}, s_i, t_{i+1}, \dots, t_n). \tag{32}$$

By assumption we have $s_i = t_i[v(\hat{w})/u(\hat{w})]$ with $(u \approx v) \in E \cup \breve{E}$. Then it is easy to see that $\sigma(t_1, \dots, t_{i-1}, s_i, t_{i+1}, \dots, t_n)$ is obtained from $\sigma(t_1, \dots, t_{i-1}, t_i, t_{i+1}, \dots, t_n)$ by replacing the subterm $u(\hat{w})$ that occurs in t_i by $v(\hat{w})$. This gives (32). Hence $\equiv_E^*$ has the substitution property.

So $\equiv_E^*$ is a congruence relation on the formula algebra. It is *substitution-invariant* in the sense that if $t(\hat{x}) \equiv_E^* s(\hat{x})$ then $t(\hat{w}) \equiv_E^* s(\hat{w})$ for any choice of terms $\hat{w}$ to subsitute for the $\hat{x}$. This is easy to see and is left as an exercise.

$E \vdash t \approx s \Longrightarrow t \equiv_E^* s$. Let $u_1 \approx v_1, \dots, u_n \approx v_n$ be a proof of $t \approx s$ from E. We prove that $u_i \equiv_E^* v_i$ for all $i \leq n$ by induction on i. If $u_i \approx v_i$ is a tautology then $u_i \equiv_E^* v_i$ because $\equiv_E^*$ is reflexive. If $u_i \approx v_i$ is a substitution instance of an equation in E, say $u_i \approx v_i$ is $q(\hat{w}) \approx r(\hat{w})$ with $q \approx r$ in E. Then $u_i \equiv_E v_i$ because v_i is obtained from u_i by replacing the subterm $q(\hat{w})$ by $r(\hat{w})$. If $u_i \approx v_i$ is obtained by an application of (*symm*), (*tran*), or (*repl*), then $u_i \equiv_E^* v_i$ because $\equiv_E^*$ is respectively symmetric, transitive, and has the substitution property.

$t \equiv_E^* s \Longrightarrow E \vdash t \approx s$. Because of the rules (*taut*), (*symm*), and (*tran*) it suffices to prove that $t \equiv_E s \Longrightarrow E \vdash t \approx s$. Suppose $t \equiv_E s$, say $s = t[v(\hat{w})/u(\hat{w})]$ where $v(\hat{x}) \approx u(\hat{x})$ is in $E \cup \breve{E}$. We prove $E \vdash t \approx s$ by the recursive depth of the principal operation symbol of $u(\hat{x})$ in the parse tree of t. If $t = u(\hat{w})$ then $s = v(\hat{w})$ and thus $t \equiv_E^* s$ by (*E-axiom*) or by (*E-axiom*) together with (*symm*). Assume $t = \sigma(q_1, \dots, q_n)$ and $u(\hat{w})$ is a subterm of q_i so that $s = \sigma(q_1, \dots, q_i[v(\hat{w})/u(\hat{w})], \dots, q_n)$. Then $q_i \equiv_E^* q_i[v(\hat{w})/u(\hat{w})]$ and hence $E \vdash q_i \approx q_i[v(\hat{w})/u(\hat{w})]$ by the induction hypothesis. Hence $E \vdash t \approx s$ by (*repl*). □

Recalling again that equations are elements of $Te_\Sigma(X)^2$ we see that $\mathrm{Cn}_\Sigma(E) = \{ \langle t, s \rangle : E \vdash t \approx s \} = \equiv_E^*$. So equational theories are exactly the subtitution-invariant congruence relations on $\mathbf{Te}_\Sigma(X)$. Then the Galois connection between consequence between Id and Mod induces a dual isomorphism between the lattice of Σ-varieties and the lattice of substution-invariant congruences on $\mathbf{Te}_\Sigma(X)$.