Studies in Systems, Decision and Control 89

Calin Belta
Boyan Yordanov
Ebru Aydin Gol

Formal Methods for Discrete-Time Dynamical Systems

Springer

$$(p, p, n, p),\ (p, p, n, p),\ (n, p, n, p),\ (n, n, n, p),\ (n, n, n, p),\ (p, n, n, p),$$
$$(p, n, n, n),\ (p, n, p, n), \ldots. \tag{1.15}$$

The word generated by the red run coincides with the above word in the first 7 entries; the 8th entry is replaced by (p, n, n, n).

If the properties of interest are formulated in terms of a set of predicates, as in Eqs. (1.9), (1.10), and (1.11), then the word of the blue run is

$$\{\pi_3\},\ \{\pi_3\},\ \{\pi_1, \pi_3\},\ \{\pi_1, \pi_2, \pi_3\},\ \{\pi_1, \pi_2, \pi_3\},\ \{\pi_2, \pi_3\},\ \{\pi_2, \pi_3, \pi_4\},\ \{\pi_2, \pi_4\}, \ldots. \tag{1.16}$$

In the red run, the 8th entry is replaced by $\{\pi_2, \pi_3, \pi_4\}$.

1.3 Simulation and Bisimulation

A number of analysis and control techniques, such as the ones presented in Part II, have been developed to handle only finite transition systems (e.g. when an explicit representation of all system states is required). In addition, these methods become computationally challenging as the size of the state set of the system increases, which limits the applicability of these methods due to the infamous "curse of dimensionality". In this section, we introduce *finite abstractions*, which can be used to reduce the size of a finite system or map an infinite system to a finite one for the purpose of analysis and control.

Intuitively, an abstraction of a transition system T preserves some of its details required for analysis and control but ignores aspects that do not influence the results.

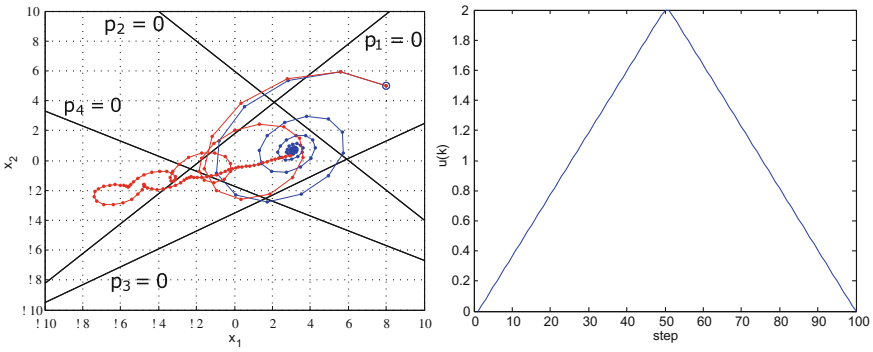


Fig. 1.8 Sample runs of the discrete-time system from Eqs. (1.13), (1.14) are shown on the *left* the *blue run* corresponds to the autonomous system ($u(k) = 0$, $k = 0, 1, 2, \ldots$), while the *red run* is produced by the input shown on the *right*. The two components of the state vector x are denoted by x_1 and x_2

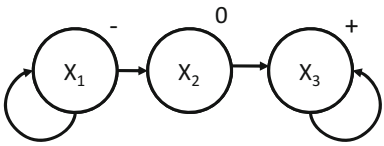


Fig. 1.9 An abstraction of the infinite transition system T with a set of states $X = \mathbb{Z}$ from Example 1.3 might preserve only the observation of a state (the sign of a number) but ignore the other details (the exact value). All states $x \in X$ with the property that $o(x) =$ "$-$" ($o(x) =$ "$+$") are equivalent and are grouped in an equivalence class X_1 (X_3)

More specifically, a state of the abstract model represents a large or infinite set of states in the original, *concrete* model that are somehow equivalent (e.g. all equivalent states might have the same observation). An abstraction could be equivalent to the concrete model with respect to the satisfaction of all specifications. Alternatively, it could provide an approximation, guaranteeing that satisfaction of a specification in the abstract model implies satisfaction in the original system. Equivalent abstractions are based on the notion of *bisimulation*, while approximate abstractions are constructed using *simulation relations*. In Sect. 4.5 we will also explore the idea of constructing abstractions which are equivalent only with respect to a given specification, and are therefore coarser than bisimulation. In all these cases, analysis or control of the large or infinite system can then be performed instead on its finite abstract model (see Fig. 1.9).

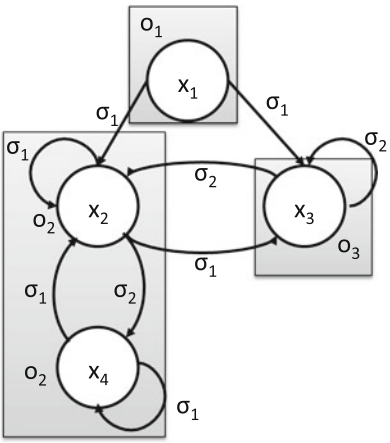
The observation map o of a transition system $T = (X, \Sigma, \delta, O, o)$ induces an equivalence relation $\sim \subseteq X \times X$ over the set of states X of T.

Definition 1.2 (*Observational equivalence*) States $x_1, x_2 \in X$ are observationally equivalent (written as $x_1 \sim x_2$) if and only if $o(x_1) = o(x_2)$.

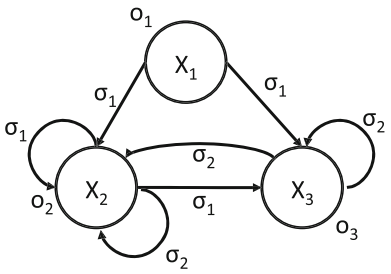
Definition 1.3 (*Quotient transition system*) The observational equivalence relation $\sim$ naturally induces a *quotient transition system* $T/_\sim = (X/_\sim, \Sigma, \delta_\sim, O, o_\sim)$, where

- the set of states $X/_\sim$ is the quotient space (i.e., the set of all equivalence classes),
- the set of inputs Σ is inherited from the original system,
- the transition relation $\delta_\sim$ is defined as follows: for states $X_i, X_j \in X/_\sim$ and input $\sigma \in \Sigma$, we include transition $X_j \in \delta_\sim(X_i, \sigma)$ if and only if there exist states x_1 and x_2 of T in equivalence classes X_i and X_j, respectively, such that x_2 is reachable from x_1 in one step under input σ (i.e., $x_2 \in \delta(x_1, \sigma)$),
- the set of observations O is inherited from T, and
- the observation $o_\sim(X)$ of a state $X_i \in X/_\sim$ is given by $o_\sim(X_i) = o(x)$ for all states x from equivalence class X_i.

Given an equivalence class $X_i \in X/_\sim$, we denote the set of all equivalent states of T in that class by $con(X_i) \subseteq X$, where *con* stands for *concretization map*. For a state X_i of $T/_\sim$ the set $con(X_i)$ is, in general, a region of T and if $\mathbb{X} \subseteq X/_\sim$ is a region of $T/_\sim$, then $con(\mathbb{X}) = \bigcup_{X_i \in \mathbb{X}} con(X_i)$ is a region of T. The observation map $o_\sim$ of $T/_\sim$ is well defined, since all states $x \in con(X_i)$ from an equivalence class $X_i \in X/_\sim$ have the same observation (i.e., $\forall x \in con(X_i), o(x) = o_\sim(X_i)$).



(a) A finite, nondeterministic control transition system T.



(b) The quotient $T/_\sim$ of control transition system T.

Fig. 1.10 The quotient $T/_\sim$ (**b**) of system T (**a**) under the observational equivalence relation $\sim$. States that are equivalent in T (i.e., states sharing the same observation) are highlighted (see Example 1.9 for additional details)

Given states $X_i, X_j \in X/_\sim$, we can assign transitions in $T/_\sim$ through computation of the successor states (Eq. (1.2)) of each equivalence class, i.e.,

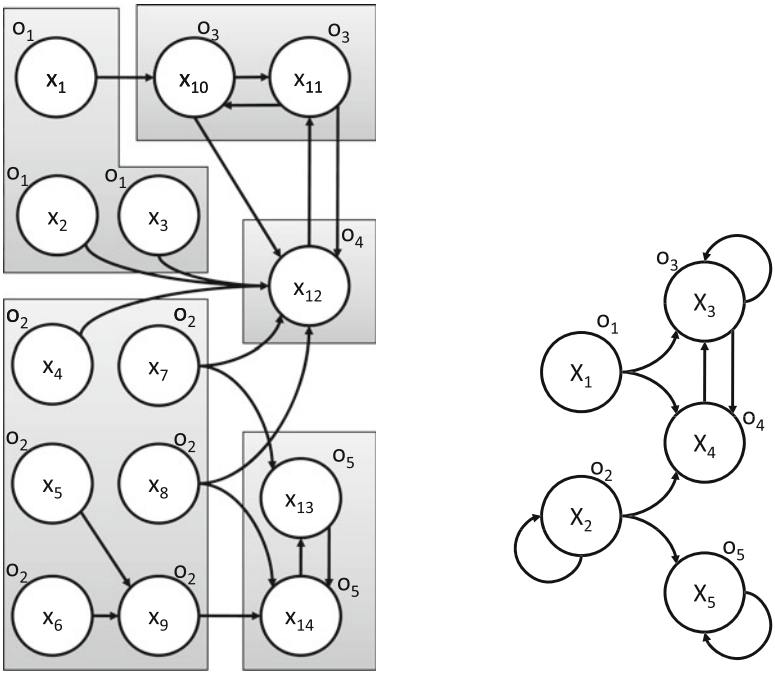
$$X_j \in \delta_\sim(X_i, \sigma) \text{ if and only if } Post_T(con(X_i), \sigma) \cap con(X_j) \neq \emptyset. \tag{1.17}$$

Equivalently, we can assign transitions in $T/_\sim$ by computing the set of predecessors (Eq. (1.3)) of each class

$$X_j \in \delta_\sim(X_i, \sigma) \text{ if and only if } con(X_i) \cap Pre_T(con(X_j), \sigma) \neq \emptyset. \tag{1.18}$$

Example 1.9 Consider the transition system from Fig. 1.1 shown again for convenience in Fig. 1.10a. The states with the same observations are equivalent. The corresponding equivalence classes and highlighted in Fig. 1.10a. The set of states of the quotient $T/_\sim$, shown in Fig. 1.10b, is $X/_\sim = \{X_1, X_2, X_3\}$. The set of all states of T in an equivalence class X_i is given by the concretization map $con()$: $con(X_1) = \{x_1\}$, $con(X_2) = \{x_2, x_4\}$, and $con(X_3) = \{x_3\}$. The observations of $T/_\sim$ are inherited from T. The transitions in $\delta_\sim$ are assigned as stated in Definition 1.3.

The above definitions can be immediately adapted for uncontrolled transition systems by considering the particular case when $|\Sigma| = 1$ followed by the deletion of the control symbol from the transition function.



(a) A finite, nondeterministic transition system T.

(b) The quotient $T/_\sim$ of transition system T.

Fig. 1.11 The quotient $T/_\sim$ (**b**) of system T (**a**) under the observational equivalence relation $\sim$. States that are equivalent in T (i.e., states sharing the same observation) are highlighted (see Example 1.10 for additional details)

Example 1.10 We construct the finite quotient $T/_\sim$ under the observational equivalence relation $\sim$ for transition system T given in Fig. 1.11a. Each state of T from the set $X = \{x_1, \ldots, x_{14}\}$ has a single observation from the set $O = \{o_1, \ldots, o_5\}$. States that have the same observation are equivalent and are highlighted in Fig. 1.11a, revealing the equivalence classes of the system. The quotient $T/_\sim$ has a set of states $X/_\sim = \{X_1, \ldots, X_5\}$ where, for $i = 1, \ldots, 5$, each state X_i represents the set of equivalent states with an observation o_i and, as a result, the observation map $o_\sim$ is clearly defined (i.e., $o_\sim(X_i) = o_i$). The set of all states of T in an equivalence class X_i is given by the concretization map $con()$: $con(X_1) = \{x_1, x_2, x_3\}$, $con(X_2) = \{x_4, \ldots, x_9\}$, $con(X_3) = \{x_{10}, x_{11}\}$, $con(X_4) = \{x_{12}\}$ and $con(X_5) = \{x_{13}, x_{14}\}$. Finally, transitions in $\delta_\sim$ are assigned as stated in Definition 1.3, resulting in the finite quotient $T/_\sim$ shown in Fig. 1.11b.

From Definition 1.3, it follows that for all states (equivalence classes) $X_i \in X/_\sim$ of $T/_\sim$, we have

$$\mathscr{L}_T(con(X_i)) \subseteq \mathscr{L}_{T/_\sim}(X_i). \tag{1.19}$$

In other words, the quotient control transition system $T/_\sim$ can produce any word w_O that can be produced by the original, concrete transition system T. However, in general there exists words in $\mathscr{L}_{T/_\sim}(X_i)$ that are *spurious* and do not represent valid behavior of T (i.e., they are not part of $\mathscr{L}_T(con(X_i))$).

Since any behavior of T can be reproduced by $T/_\sim$, we say that $T/_\sim$ *simulates* T. As we will discuss later in the book, this guarantees that, if a linear temporal logic formula is satisfied at some state X_i of $T/_\sim$, then the formula will be satisfied at all the states of T contained in $con(X_i)$.

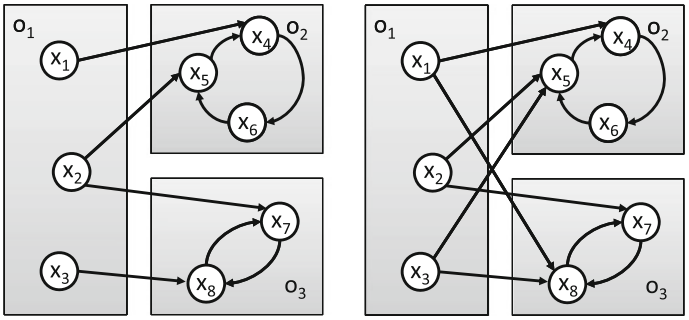
Definition 1.4 (*Bisimulation*) The equivalence relation $\sim$ induced by the observation map o is a bisimulation of a transition system $T = (X, \Sigma, \delta, O, o)$ if, for all states $x_1, x_2 \in X$, and all inputs $\sigma \in \Sigma$, if $x_1 \sim x_2$ and $x_1' \in \delta(x_1, \sigma)$, then there exist $x_2' \in X$ such that $x_2' \in \delta(x_2, \sigma)$ and $x_1' \sim x_2'$.

If $\sim$ is a bisimulation, then the quotient transition system $T/_\sim$ is called a *bisimulation quotient* of T, and the transition systems T and $T/_\sim$ are called *bisimilar*. To explicitly distinguish between a simulation and a bisimulation, we sometimes denote the latter by $\approx$ and use $T/_\approx$ to denote a bisimulation quotient. In other words, the quotient $T/_\approx$ is the quotient $T/_\sim$ when Definition 1.4 is satisfied.

Simulations and bisimulations relations are generally defined between transition systems sharing the same sets of observations, through equivalence relations between their states. In this book, we restrict our attention to equivalence relations defined over the states of a transition system, and the simulation/bisimulation relations are between the original system and its quotient. These notions are usually called observational simulation/bisimulation but throughout the rest of this book we will simply denote them as simulation/bisimulation.

For example, in Fig. 1.10b, the quotient $T/_\sim$ is not a bisimulation of the transition system T shown in Fig. 1.10a. Note that x_2, which is in the equivalence class X_2 has a transition to some state in the equivalence class X_3 under input σ_1. However, x_4, which is also in the equivalence class X_2 does not have a transition to some state in the equivalence class X_3 under σ_1. Examples of transition systems with no inputs for which the observational equivalence relation $\sim$ is not or is a bisimulation are given in Fig. 1.12a, b, respectively.

Definition 1.4 establishes bisimulation as a property of the quotient $T/_\sim$, when transitions originating at equivalent states in T satisfy certain conditions. In the following, for the particular case of transition systems with no inputs, we consider alternative conditions guaranteeing that the quotient $T/_\sim$ is a bisimulation quotient, which are easier to test computationally.



(a) $\sim$ is not a bisimulation. (b) $\sim$ is a bisimulation.

Fig. 1.12 The observational equivalence relation $\sim$ (shown as *shaded rectangles*) is a bisimulation for the transition system from (**b**) but not for the one from (**a**) (for both systems only the observations for the entire equivalence classes are given)

Proposition 1.1 *For a transition system $T = (X, \delta, O, o)$, the equivalence relation $\sim$ is a bisimulation if the quotient $T/_\sim$ is deterministic.*

Proof Assume by contradiction that $\sim$ is not a bisimulation. Then, there exist $X_i, X_j \in X/_\sim, x_1, x_2 \in con(X_i)$, and $x_1' \in con(X_j)$ such that $x_1 \to x_1'$ but there does not exist $x_2' \in con(X_j)$ such that $x_2 \to x_2'$. However, since T is nonblocking, there exists $x_2'' \in X$ and $X_k \in X/_\sim$, $X_k \neq X_j$ such that $x_2'' \in \delta(x_2)$, where $x_2'' \in con(X_k)$. In the quotient $T_\sim$, this induces transitions $X_j \in \delta_\sim(X_i)$ and $X_k \in \delta_\sim(X_i)$, which implies that $T/_\sim$ is nondeterministic and contradicts the hypothesis. ■

Proposition 1.1 offers a computationally attractive sufficient condition for bisimulation for transition systems with no inputs, where only the number of outgoing transitions from each state in the quotient is counted. For deterministic transition systems, this result becomes stronger:

Proposition 1.2 *An equivalence relation $\sim$ defined on a deterministic transition system $T = (X, \delta, O, o)$ is a bisimulation if and only if the quotient $T/_\sim$ is deterministic.*

Proof From Proposition 1.1 it follows that if the quotient is deterministic then the equivalence relation is a bisimulation. Assume by contradiction that $T/_\sim$ is not deterministic. Then, there exist $X_i, X_j, X_k \in X/_\sim$ such that $X_j \in \delta_\sim(X_i)$ and $X_k \in \delta_\sim(X_i)$. However, since $\sim$ is a bisimulation, there exists $x_i, x_j, x_k \in X$, $x_i \in con(X_i)$, $x_j \in con(X_j)$, $x_k \in con(X_k)$ such that $x_j \in \delta(x_i)$ and $x_k \in \delta(x_i)$, which implies that T is non-deterministic and contradicts the hypothesis. ■

Relevant to model checking and analysis problems that we will define later in the book, as an immediate consequence of bisimulation, we can guarantee the language equivalence between the quotient $T/_\approx$ and the concrete system T. In other words, for all states $X_i \in X/_\approx$, it holds that

$$\mathscr{L}_T(con(X_i)) = \mathscr{L}_{T/\approx}(X_i). \tag{1.20}$$

In Definition 1.4, we gave conditions on the transitions originating at equivalent states in system T, required for the observational equivalence relation $\sim$ to be a bisimulation. Following from Definition 1.4, a computationally attractive characterization of bisimulation can be given by considering the set of predecessors (Eq. (1.5)) of each equivalence class.

Theorem 1.1 (Bisimulation characterization) *The equivalence relation $\sim$ is a bisimulation for a transition system $T = (X, \Sigma, \delta, O, o)$ if and only if for all equivalence classes $X_i \in X/_\sim$ and for all inputs $\sigma \in \Sigma$, $Pre_T(con(X_i), \sigma)$ is either empty or a finite union of equivalence classes. Equivalently, the bisimulation property from Definition 1.4 is violated at state $X_i \in X/_\sim$ if there exist an input $\sigma \in \Sigma$ and a state $X_j \in X/_\sim$, such that*

$$\emptyset \subset con(X_i) \cap Pre_T(con(X_j), \sigma) \subset con(X_i). \tag{1.21}$$

As a particular case, for a transition system $T = (X, \delta, O, o)$ with no inputs, the bisimulation characterization takes the form of Eq. 1.21 with σ removed.

Example 1.11 Consider the quotient $T/_\sim$ (Fig. 1.13) of transition system T from Fig. 1.12a. Using the characterization from Theorem 1.1, we can verify that the bisimulation property (Definition 1.4) is violated by equivalence class $X_1 \in X/_\sim$. The set $Pre(con(X_3)) = \{x_2, x_3\}$ is not empty and has a nonempty intersection with $con(X_1) = \{x_1, x_2, x_3\}$ but there exist a state $x_1 \in con(X_1)$ such that $x_1 \notin Pre((con(X_3))$. Equivalently, $Pre((con(X_3))$ is not a finite union of equivalence classes but is, in fact, a subset of an equivalence class. Since the characterization from Theorem 1.1 is violated at state X_1 (i.e., Eq. (1.21) is satisfied at that state), the quotient $X/_\sim$ is not bisimilar with T.

Algorithm 1 $\approx$=BISIMULATION(T): Construct the coarsest observation-preserving bisimulation quotient $\approx$ of $T = (X, \Sigma, \delta, O, o)$

1: Initialize $\sim_r := \sim$
2: **while** there exist equivalence classes $X_i, X_j \in X/_{\sim_r}$ and $\sigma \in \Sigma$ such that
$\emptyset \subset con(X_i) \cap Pre_T(con(X_j), \sigma) \subset con(X_i)$ **do**
3: Construct equivalence class X_1 such that $con(X_1) := con(X_i) \cap Pre_T(con(X_j), \sigma)$
4: Construct equivalence class X_2 such that $con(X_2) := con(X_i) \setminus Pre_T(con(X_j), \sigma)$
5: $X/_{\sim_r} := X/_{\sim_r} \setminus \{X_i\} \bigcup \{X_1, X_2\}$
6: **end while**
7: return $\sim_r$ ($\sim_r$=$\approx$)

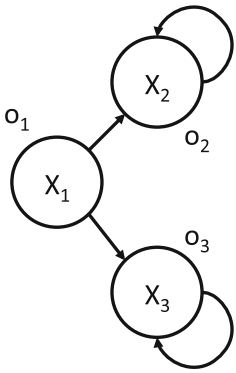


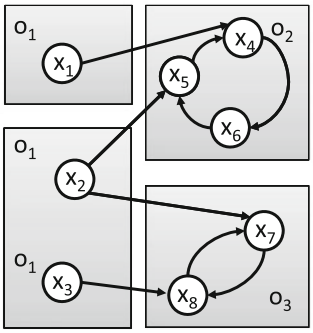
Fig. 1.13 Quotient $T/_{\sim}$ is the same for the transition systems from Fig. 1.12a, b. While $T/_{\sim}$ is a bisimulation quotient for the system from Fig. 1.12b, it only simulates the one from Fig. 1.12a. For that system, the bisimulation characterization from Theorem 1.1 is violated at state X_1. See Example 1.11 for details

Equation (1.21) leads to an approach for the construction of the coarsest bisimulation $\approx$. Given a transition system T, the iterative procedure known as the "bisimulation algorithm" (summarized as Algorithm 1) starts with the observational equivalence relation $\sim$ and uses it to identify equivalence classes from $X/_{\sim}$ that satisfy Eq. (1.21). Then, it iteratively refines these classes until the characterization from Theorem 1.1 is satisfied. This guarantees that the equivalence relation $\sim_r$ returned after the algorithm terminates is indeed $\approx$—a bisimulation of T, which can be used to construct the bisimulation quotient $T/_{\approx}$.

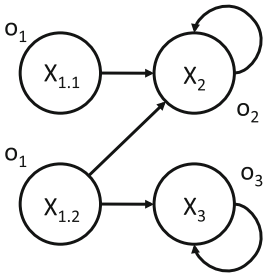
Example 1.12 Consider transition system T from Fig. 1.12a, with quotient $T/_{\sim}$ (Fig. 1.13) induced by the observational equivalence relation $\sim$. As already discussed in Example 1.11, $T/_{\sim}$ is not a bisimulation quotient (the characterization from Theorem 1.1 is violated at state X_1, for which Eq. (1.21) is satisfied). In order to obtain the coarsest observation preserving equivalence relation of T, the bisimulation algorithm (Algorithm 1) is applied to T.

First, equivalence classes X_1 and X_3 of $T/_{\sim}$ are considered, where $\emptyset \subset con(X_1) \cap Pre(con(X_3)) \subset con(X_1)$. Equivalence class X_1 is refined into $X_{1.1}$ and $X_{1.2}$, such that $con(X_{1.1}) = con(X_1) \setminus Pre(con(X_3)) = \{x_1\}$ and $con(X_{1.2}) = Pre(con(X_3)) \cap con(X_1) = \{x_2, x_3\}$, which leads to the construction of the intermediate equivalence relation $\sim_1$ represented in Fig. 1.14a. Equivalence relation $\sim_1$ induces quotient $T/_{\sim_1}$ (Fig. 1.14b) but is still not a bisimulation (the characterization from Theorem 1.1 is violated for $X_{1.2} \in X/_{\sim_1}$). Therefore, Algorithm 1 is applied again.

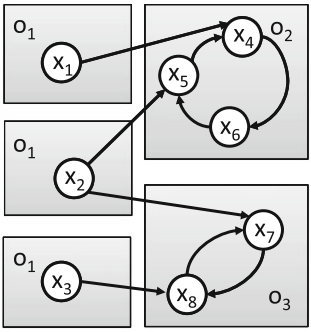
Equivalence classes $X_{1.2}$ and X_2 of $X/_{\sim_1}$ are then considered and $X_{1.2}$ is refined into $X_{1.2.1}$ and $X_{1.2.2}$, such that $con(X_{1.2.1}) = con(X_{1.2}) \cap Pre(con(X_2)) = \{x_3\}$ and $con(X_{1.2.2}) = con(X_{1.2}) \setminus Pre(con(X_2)) = \{x_2\}$. The resulting equivalence relation $\sim_2$ represented in Fig. 1.14c is a bisimulation and induces the bisimulation quotient $T/_{\sim_2} = T/_{\approx}$ shown in Fig. 1.14d.



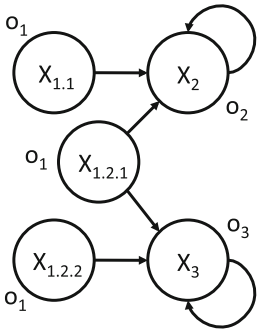
(a) Equivalence classes $X/_{\sim_1}$ after refinement of $X_1 \in X/_{\sim}$.



(b) Quotient $T/_{\sim_1}$.



(c) Equivalence classes of $T/_{\sim_2}$ after refinement of $X_{1.2} \in X/_{\sim_1}$



(d) Quotient $T/_{\sim_2}$.

Fig. 1.14 The equivalence relations $\sim_1$ and $\sim_2$, obtained at successive steps while applying the bisimulation algorithm to system T from Fig. 1.12a, induce quotients $T/_{\sim_1}$ and $T/_{\sim_2}$, respectively. See Example 1.12 for additional details

1.4 Notes

In this chapter we introduced transition systems as a modeling formalism for a wide range of processes [11]. Besides capturing the behavior of many processes directly, transition systems provide a semantical model for various high-level formalisms for concurrent systems including Kripke structures [45], process algebras [88, 131, 132], statecharts [81] and Petri nets [140] (also see [39] for additional discussion on discrete event systems). In addition, the formalisms of Mealy and Moore machines, which are related to finite state automata and are commonly used in hardware synthesis and analysis, can be described by transition systems [110]. Finally, in biological applications, Boolean networks [101] are a popular model approximating the behavior of large genetic and signaling networks. The state spaces of both Boolean networks and their extending qualitative networks [154] and generalized logical networks [166] can also be expressed as finite, and in many cases deterministic, transition systems.

We showed that infinite transition systems are rich enough to describe dynamical (control) systems. Such embeddings have been proposed by several authors [5, 7, 72, 74, 105, 106, 138, 162, 163, 180, 184, 185]. While such embeddings can be easily defined for continuous-time systems as well [105, 106, 138], we do not give these definitions as the focus in this book is on discrete-time systems only.

The transition system definition that we consider here is somewhat different from the ones encountered elsewhere [11, 15, 45]. We allow a state $x \in X$ of the system to have only a single observation from the set of observations O, which is given by the observation map $o(x) \in O$. In more classical definitions, a state can have several observations from the set O (or, equivalently, a state might satisfy several propositions from a given set) and, as a result, the observation map o gives the set of observations $o(x) \in 2^O$ at state $x \in X$. Such a formulation can be reduced without loss of generality to the one we use by redefining the set of observations (i.e., by defining a new observation for each subset from 2^O). As it will become clear in Chap. 2, this assumption will also induce a slightly different interpretation of the semantics of the temporal logic formulas.

In this book, we deal with transition systems that are not initialized (i.e., we do not specify a set of initial states for the system) unlike the systems that are commonly encountered elsewhere [15, 45]. Such a definition is more appropriate for the analysis applications we consider in subsequent chapters. We still formulate the model checking problem as deciding if all runs from a given region satisfy a specification, which is equivalent to model checking a transition system that is initialized at that region. In other texts, only transition systems that have been initialized at a single state can be deterministic but according to our formulation, determinism is a property of the transitions and not of the initial states of a system. This allows us to search for sets of initial states rather than individual initial states as part of the analysis problem we consider in Chap. 4, even when deterministic systems are considered.

As probabilistic dynamics are not covered in this book, the transition systems defined in this chapter capture only purely deterministic and nondeterministic behaviors. The probabilistic version of the transition system considered here is the well known Markov decision process (MDP), in which the inputs enable transitions with given probability distributions among the states of the system [15]. Throughout the book, we also assume that the states of the system are observable. In other words, the current state of the system is known and available for state-feedback control. Readers interested in systems for which this assumption is relaxed are referred to transition systems with nondeterministic observations [8], for which the current state is known only to belong to a given set, and probabilistic versions such as hidden Markov models (HMM) [26], partially observable Markov decision processes (POMDP) [98], and mixed observability Markov decision processes (MOMDPs) [135].

The notions of simulation and bisimulation that we consider here are relations between systems and their quotients. Such relations can be defined, in general, between systems sharing the same sets of observations [11]. There also exist relaxed notions of bisimulations, such as alternating [6], weak [131], probabilistic [116], and approximate bisimulations [67, 145], which go beyond the scope of this book.