On supervisor reduction in discrete-event systems

A. F. VAZ† and W. M. WONHAM†

A supervisory controller (supervisor) $S$ for a discrete-event system can be modelled on a recognizer $R$ for the language corresponding to the supervisory task to be accomplished. It is shown that simpler 'reduced' supervisors can be constructed by the use of covers of the state set of $R$ and that any mildly restricted supervisor is a reduction of $R$ in this sense. The reduction procedure is time-exponential with respect to the size of the state set of $R$.

I. Introduction

Discrete-event systems (DES) can be described as systems or processes that are discrete, asynchronous and possibly nondeterministic. Examples of systems of this type include manufacturing systems, communication protocols and the like. A control theory for a general class of DES has been proposed in Ramadge and Wonham (1983); see also the thesis by Ramadge (1983), the summary articles by Ramadge and Wonham (1984), Wonham and Ramadge (1984 a, b) and Wonham (1985 b) and the detailed reports by Ramadge and Wonham (1983), Wonham and Ramadge (1983 a, b), Wonham (1985 a) and Lin and Wonham (1985). The reader is referred to these sources for general background and motivation, as well as for details of the mathematical setup, on which we shall draw freely.

In this theory the DES to be controlled is modelled as a tuple

$$
G = (\Sigma, \Sigma_\text{u}, \Sigma_\text{s}, Q, \delta, q_0, Q_m)
$$

where

$$
\Sigma = \Sigma_\text{u} \cup \Sigma_\text{s}
$$

is an alphabet of event labels consisting of uncontrollable elements ($\Sigma_\text{u}$) and controllable elements ($\Sigma_\text{s}$); $Q$ is the state set; $\delta: \Sigma \times Q \to Q$ is the state transition function (a partial function, defined for each $q \in Q$ on a subset of the $s \in \Sigma$); $q_0 \in Q$ is the initial state; and $Q_m \subset Q$ is the subset of 'marked' states. Let $\Sigma^*$ denote the set of all finite strings $s$ of elements of $\Sigma$, including the empty string $\epsilon$. By finite iteration of $\delta$ starting from $q_0$, $G$ is established as the generator of a formal language

$$
L(G) = \{ s \in \Sigma^* \mid \delta(s, q_0) \text{ is defined} \}
$$

while

$$
L_m(G) = \{ s \in \Sigma^* \mid \delta(s, q_0) \in Q_m \}
$$

is the subset of $L(G)$ of 'marked' strings representing 'completed tasks' of the DES that $G$ is intended to model.

Received 18 November 1985.
† Systems Control Group, Department of Electrical Engineering, University of Toronto, Toronto, Ontario, Canada M5S 1A4.
A supervisor for $G$ is a pair $S = (S, \psi)$, where

$$S = (\Sigma, X, \xi, x_0, X_m)$$

is an automaton with state set $X$, initial state $x_0$, a marked subset $X_m \subseteq X$, and (partial) transition function $\xi : \Sigma \times X \to X$; while

$$\psi : \Sigma \times X \to \{0, 1, d\c$$

is a control law such that

$$\begin{align*}
\psi(\sigma, x) &= 1, \\
\sigma &\in \Sigma_x, x \in X \\
\psi(\sigma, x) &= 0, 1, d, \\
\sigma &\in \Sigma_x, x \in X
\end{align*}$$

$S$ is considered to be driven externally by the stream of output elements $\sigma$ of $G$; while in turn, with $S$ in state $x$, the transitions (labelled) $\sigma$ of $G$ are subject to the control $\psi(\sigma, x)$. If $\psi(\sigma, x) = 0$ then $\sigma$ is 'disabled' (prohibited from occurring), while if $\psi(\sigma, x) = 1$ then $\sigma$ is 'enabled' (permitted but not forced to occur). The notation $d$ ('don't care') means that assignment to $(\sigma, x)$ of either 0 or 1 would make no difference to the behaviour of $G$. The transition function $\delta : \Sigma \times Q \to Q$ of $G$ is then replaced by

$$\delta_{\psi} : \Sigma \times X \to Q$$

where

$$\delta_{\psi}(\sigma, x, q) = (\delta(\sigma, x), q)$$

if $\psi(\sigma, x) = 1$ and is undefined otherwise. In this way one obtains a closed-loop feedback structure $S/G$ (as supervised by $S$). $S/G$ is thereby another DES:

$$S/G = (\Sigma, \Sigma_x, \Sigma_{\psi}, X \times Q, \xi \times \delta_{\psi}(x_0, q_0), X_m \times Q_m)$$

with (partial) transition function defined by

$$(\xi \times \delta_{\psi})(\sigma, x, q) = (\xi(\sigma, x), \delta_{\psi}(\sigma, x, q))$$

whenever both components on the right are defined. We denote the corresponding closed-loop languages by $L(S/G)$, $L_{\psi}(S/G)$. For simplicity we temporarily confine attention to $L(S/G)$.

It was shown in Ramadge and Wonham (1983) that, if $K \in \Sigma^*$ is a language such that there exists a supervisor $S = (S, \psi)$ with

$$K = L(S/G),$$

then $S$ can be modelled on a recognizer for $K$. In practice the transition structure of $K$ is often considerably more complex than is necessary for $S$ to perform its role in $S$. It is then of interest to reduce $S$ to a simpler supervisor, say $S$, such that control action is preserved, namely

$$K = L(S/G).$$

In Ramadge and Wonham (1983) such a reduction was achieved by a suitable partition of the state set $X$ of $S$. The objectives of the present article are (i) to generalize this construction by the use of 'covers' instead of partitions; (ii) to obtain generalizes versions of the Quotient Structure Theorem of Ramadge and Wonham (1983); and (iii) to present an algorithm by which a minimal (i.e. 'optimal') state structure for $S$ can be computed.

2. Main structural results

For a partial function (e.g. $\zeta$) we write $\zeta(\sigma, x)$ to mean that $\zeta(\sigma, x)$ is defined. Also, if $K \in \Sigma^*$ is a language and $s, t \in \Sigma^*$ are arbitrary strings, then $s \equiv t \mod K$

means that $s,t$ belong to the same Nerode equivalence class of $K$, namely for all $w \in \Sigma^*$, $sw \in K$ if and only if $tw \in K$. If $K$ is regular, $|K|$ denotes the number of $K$-equivalence classes of $\Sigma^*$. Finally, we refer to Ramadge and Wonham (1982, 1983) for the notion of controllability of a language.

2.1. Standard supervisor

Write $K := L(S/G), L := L(G)$. Following Ramadge and Wonham (1983), we say that a supervisor $S = (S, \psi)$ for $G$ is complete if, for all $s \in K$, the two conditions $s \in L$ and $\psi(s, x) = 1$ with $x = \zeta(s, x_0)$ together imply $\psi(s, x)$, namely $s \in L(S/G)$. Also, a supervisor $S = (S, \psi)$ will be called normal if the control law $\psi$ is defined with as much flexibility as possible, namely enablement (1) or disablement (0) is assigned to a pair $(\sigma, x), \sigma \in \Sigma_x, x \in X$, only when necessary. Precisely, $S$ is normal provided for all $x \in X$

$$\begin{align*}
\psi(\sigma, x) &= 0 \text{ if there is } s \in K \text{ such that } \zeta(s, x_0) = x, s \in L, \text{ and } s \notin K \\
\psi(\sigma, x) &= 1 \text{ if there is } s \in K \text{ such that } \zeta(s, x_0) = x, \text{ and } s \in K
\end{align*}$$

Thirdly, $S$ will be called strongly $K$-accessible if for every $x \in X$ there exists $s \in K$ such that $\zeta(s, x_0) = x$ (i.e. $S$ is $K$-accessible) and, for all $\sigma \in \Sigma$ and $x \in X$, $\zeta(\sigma, x)$ only if there is $s \in K$ such that $\zeta(s, x_0) = x$ and $s \notin K$. Finally we shall say that a supervisor $S$ is standard if it is complete, normal and strongly $K$-accessible. It is straightforward to verify that any complete supervisor may be replaced by a standard version without changing the control action with respect to the behaviour of $G$. From now on it will be a tacit assumption that $S$ is standard.

2.2. Covers and supervisor reduction

To develop the main results we define a cover of $S$ to be a family $C = \{X_j, i \in I\}$ of subsets of $X$ with the following properties (cf. Zeiger 1968):

$$\begin{align*}
&\forall i X_i \neq \emptyset; \\
&\text{for a subset } I_m \subseteq I, \\
&X_m = \cup \{X_i | i \in I_m\}, \quad X - X_m = \cup \{X_i | i \in I - I_m\};
\end{align*}$$

$$\begin{align*}
&\forall i, \sigma) (\exists y \in X_i) \exists (\sigma, y) \\
&(\forall i, \sigma) (\exists y \in X_i) \exists (\sigma, y) \\
&\Rightarrow (\exists i) (\forall x \in X_i) \exists (\sigma, x) \\
&\text{[for brevity we write this property as]} \\
&\forall i, \sigma) (\exists y \in X_i) \exists (\sigma, x) \\
&\forall i, \sigma) (\exists y \in X_i) \exists (\sigma, x) \\
&\text{Thus a cover of } S = (S, \psi) \text{ is simply a covering of the state set } X \text{ by non-empty subsets, such that the marked states } X_m \text{ are covered separately from the unmarked states } (X - X_m), \text{ the cover elements behave consistently under the action of the transition function } \xi, \text{ and a cover element exhibits uniform control action at those states where control matters.}
\end{align*}$$

Fix $i \in I, \sigma \in \Sigma$. If $\zeta(\sigma, x) \neq \emptyset$ for some $x \in X_{\sigma}$, select $j \in I$ such that $\zeta(\sigma, x) \in X_j$ for all
such $x$. A triple $(i, \sigma, j)$ as just described will be called a cover triple. Also, write
\[ \xi^{\sigma}: X \to X: x \mapsto (\xi(x), \sigma, j) \] (pfn)
for the $\sigma$-section of $\xi^\sigma$, etc., and
\[ \psi^\sigma_i: X_i \to X \]
for the restrictions of $\xi^\sigma_i, \psi^\sigma$ to the cover elements $X_i, X_j$, as indicated. The cover can then be displayed as follows, where $(i, \sigma, j)$ ranges over all cover triples:

![Diagram](image_url)

Figure 1.

Here the vertical arrows are subset insertions, and the left diagram commutes whenever $\xi^\sigma_i j$ is defined. The cover thus provides a set of 'local' descriptions of the 'global' functions $\xi^\sigma, \psi^\sigma$.

The case of interest is $|C| < |X|$, where $| \cdot |$ denotes the number of elements (finite or denumerable) of its argument. Given $C$ with this property, we define a reduced supervisor $S = (\bar{S}, \bar{\psi})$ as follows:

1. Select $i_0 \in I$ such that $x_{i_0} \notin X_{i_0}$.
2. Define $\xi: \Sigma \times I \to I$ (pfn) as follows:
   - For $\sigma \in \Sigma, i \in I$ select $j \in I$ such that $(i, \sigma, j)$ is a cover triple, and let $\xi(\sigma, i) := j$.
3. Define $\psi: \Sigma \times I \to \{0, 1, dc\}$ as follows:
   - For $\sigma \in \Sigma, i \in I, j \in I$ if there exists $x \in X_j$ such that $\psi(\sigma, x) \neq dc$ then let $\psi(\sigma, x) := \psi(\sigma, x)$;
   - Otherwise let $\psi(\sigma, x) := dc$.

We shall say that $\bar{S}$ is based on $C$. In general the reduced supervisor is determined by selections that in part are arbitrary, but this fact is of no account in what follows.

We wish to show that $\bar{S}$ 'tracks' any input string $s \in \Sigma^*$ that is defined for $S$. For this we first define the extension
\[ \bar{\xi}: \Sigma^* \times X \to X \] (pfn)
as follows: for $x \in X$, $\bar{\xi}(1, x) := x$, and $\bar{\xi}(\sigma, x) := (\bar{\xi}(s(\sigma, x)), \bar{\xi}(\sigma, x))$ for $s \in \Sigma^*, \sigma \in \Sigma$; namely $\bar{\xi}(s, x) \neq 1$ just when $x' := \bar{\xi}(s, x') \neq \bar{\xi}(s, x)$. Similarly,
\[ \bar{\psi}: \Sigma^* \times I \to I \] (pfn)
is the extension of $\bar{\psi}$ as introduced above.

\[ \textbf{Lemma 1} \]
If $x \in X_j, s \in \Sigma^*$ and $\bar{\xi}(s, x) \neq 1$ then $\bar{\xi}(s, j); \text{ furthermore } \bar{\xi}(s, i) = j \text{ implies } \bar{\xi}(s, x) \in X_j$

\[ \textbf{Proof} \]
We use induction on the length $|s|$ of $s$. The result is clear for $|s| = 0$, namely $s = 1$. Assuming the result for $|s| = n$, suppose that $\bar{\xi}(s(\sigma, x))$ then $\bar{\xi}(s, x)$; so that $\bar{\xi}(s, x)$. Let $k = \bar{\xi}(s, i)$, $j = \bar{\xi}(s, x)$. Then $y \in X_j$ and $\bar{\xi}(s, i) = \bar{\xi}(\sigma, (s, i)) = \bar{\xi}(\sigma, y)$.

$\bar{\xi}(s(\sigma, x)) = \bar{\xi}(\sigma, (s, x)) = \bar{\xi}(\sigma, y)$

By definition of $\xi^\sigma$, we have $\bar{\xi}(s, k) = j \text{ implies } \bar{\xi}(s, y) \in X_j$, and the result follows for $|s| = n + 1$.

\[ \textbf{Corollary} \]
If $x := \bar{\xi}(s, x)$ and $\psi(\sigma, x) \neq dc$ then
\[ \psi(\sigma, x) = \psi(\bar{\xi}(s, x)), \text{ with } j = \bar{\xi}(s, i_0) \]

\[ \textbf{Proof} \]
By Lemma 1, $x \in X_j$, and the result follows by definition of $\bar{\psi}$.

We can now prove our first main result.

\[ \textbf{Theorem 1} \]
Let $S = (\bar{S}, \bar{\psi})$ be a standard supervisor for $G$, let $C = \{X_i, i \in I\}$ be a cover for $S$, and let $\bar{S} = (\bar{S}, \bar{\psi})$ be a reduced supervisor based on $C$. Then
1. $\bar{S}(G) \subseteq S(G)$
2. $\bar{S}(G) = S(G)$
3. $\bar{S}$ is complete.

\[ \textbf{Proof} \]
By induction on the length $|s|$ of strings $s$. Clearly 1 is true. Assume $s \in \bar{S}(G) \cap S(G)$. Assume $s = s(\sigma, q_0, i, q_1)$ in $\bar{S}(G)$ and let $x = \bar{\xi}(s, x), j = \bar{\xi}(s, j)$ and $q_1 = \bar{\xi}(s, q_1)$. Suppose $s \in \bar{S}(G)$, so that $\bar{\xi}(s, x), \bar{\xi}(s, q_1)$ and $\bar{\psi}(\sigma, x) \in \bar{\Sigma}(G)$. Conversely, if $s \in \bar{S}(G)$ then $\bar{\xi}(s, q_1)$. Similarly, if $s \in \bar{S}(G)$ then $\bar{\xi}(s, q_1)$. Thus $s \in \bar{S}(G)$ if and only if $s \in S(G)$. Conversely, if $s \in \bar{S}(G)$ then $\bar{\xi}(s, x) \in \bar{S}(G)$ and $\bar{\xi}(s, x) \in X_j$. By definition of $\bar{\psi}$, $\bar{\xi}(s, j) \neq 1$ implies $\bar{\xi}(s, x) \in X_j$. By definition of cover, $X_j \subseteq X_j$; i.e., $j \in I$ so $s \in \bar{S}(G)$.
Conversely, if \( s \in L_s(\mathbf{S}, \mathbf{G}) \) then \( \xi(s, x_0) \in \mathbf{Q}_s \) and \( \hat{\xi}(s, x_0) \in \mathbf{I}_s \). Since \( L_{\text{red}}(\mathbf{S}, \mathbf{G}) \subset \mathbf{L}(\mathbf{S}, \mathbf{G}) \) we have \( \hat{\xi}(s, x_0) \) and
\[
(\xi, x_0) \in \mathbf{I}_s(\mathbf{S}, \mathbf{G}) \quad (\text{i})
\]
\[
(\xi, x_0) \in \mathbf{X}_s(\mathbf{S}, \mathbf{G}) \quad (\text{ii})
\]
\[
(\xi, x_0) \in \mathbf{Y}_s(\mathbf{S}, \mathbf{G}) \quad (\text{iii})
\]
\[
(\xi, x_0) \in \mathbf{Z}_s(\mathbf{S}, \mathbf{G}) \quad (\text{iv})
\]

While our construction does not guarantee that \( \mathbf{S} \) is normal or strongly \( K \)-accessible, the fact that \( \mathbf{S} \) is complete ensures that it can be modified to a standard version if desired. For this it suffices to eliminate any \( K \)-inaccessible states in \( \mathbf{S} \) and redundant transitions in the definition of \( \hat{\xi} \), and to weaken the definition of \( \hat{\xi} \) wherever possible, from 0 or 1 to \( dc \).

A final observation is of interest. Suppose \( C = \{ X_i \} \) is a cover of \( \mathbf{S} \) such that \( X_i \subseteq X_j \) for some \( i \neq j \). Consider a supervisor \( \mathbf{S'} \) based on \( C \). Then \( \mathbf{S} \) can be simplified by merging its state \( i \) with \( j \). For this it suffices to redefine (if necessary) \( \hat{\xi}(i, j) = \xi(i, j) \), and to note that \( \hat{\xi}(j, j) = \hat{\xi}(i, i) \). In the transition graph of \( \mathbf{S} \) the node labelled \( i \) can then be glued on to the node labelled \( j \) with no change in closed-loop behaviour, namely the languages \( L(\mathbf{S'}, \mathbf{G}) \) remain the same.

2.3. First generalized quotient structure theorem

Before developing the first generalized quotient structure theorem, we must introduce one more technical definition relating to supervisor structure. We first recall the definition Ramadge and Wonham (1983) that \( \mathbf{S} \) is \( \mathbf{L}(\mathbf{S}, \mathbf{G}) \)-reduced if, whenever \( s, s' \in L(\mathbf{S}, \mathbf{G}) \) are equivalent mod \( L(\mathbf{S}, \mathbf{G}) \), then \( \xi(s, x_0) = \xi(s', x_0) \). In other words, strings that belong to the same Nerode class of \( L(\mathbf{S}, \mathbf{G}) \) must lead to the same state of the automaton \( \mathbf{S} \) of \( \mathbf{S} \). This condition on \( \mathbf{S} \) is somewhat restrictive, and we shall relax it as follows. \( \mathbf{S} \) will be called weakly \( \mathbf{L}(\mathbf{S}, \mathbf{G}) \)-reduced if, whenever \( s, s' \in L(\mathbf{S}, \mathbf{G}) \) are equivalent mod \( L(\mathbf{S}, \mathbf{G}) \), and for some \( \sigma \in L(\mathbf{S}) \), \( \xi(s, x_0) = \xi(s', x_0) \). In other words, the weakened condition is satisfied if the condition of \( \mathbf{L}(\mathbf{S}, \mathbf{G}) \)-reduction is restricted to those Nerode classes of \( L(\mathbf{S}, \mathbf{G}) \) with the property that some string \( \sigma \) of the class has an extension \( \sigma' \) that belongs to \( L(\mathbf{S}, \mathbf{G}) \) although not to \( L(\mathbf{S}, \mathbf{G}) \).

We next bring in an obvious definition of supervisor isomorphism. Let \( \mathbf{S} \) and \( \mathbf{S'} \) be supervisors for \( \mathbf{G} \); the notations \( \mathbf{S} \) will be used to label states of \( \mathbf{S} \). Then \( \mathbf{S} \) and \( \mathbf{S'} \) are isomorphic if there is a bijection \( \theta : X \to \mathbf{X} \) such that \( \theta(x_0) = x_0 \) and \( \theta(x_0) = X_\theta \). Two states \( \sigma, x \) such that the following evaluations are defined:
\[
\hat{\xi}(\sigma, x) = \hat{\xi}(\theta(\sigma), x)
\]
Our final preliminary definition is that of a recognizer. Let \( M \subset \Sigma^* \) be a language and \( K \) denote the prefix closure of \( M \). An automaton
\[
\hat{\mathbf{S}} = (\Sigma, \hat{X}, \hat{\xi}, \hat{x}_0, \hat{X}_w)
\]
is a recognizer for \( M \) if
\[
(\text{i}) \quad K = \{ s \in \Sigma^* | \hat{\xi}(s, x_0) \}
\]
\[
(\text{ii}) \quad \hat{\mathbf{S}} \quad \mathbf{S} \quad \text{is a recognizer for } M \text{ if and only if}
\]
\[
(\text{iii}) \quad \mathbf{S} \quad \text{is a recognizer for } M \text{ if and only if}
\]
Notice that a recognizer is strongly \( K \)-accessible. We shall only be concerned with the case where \( M \) is closed, i.e. \( K = M \).

We can now state our second main result. Assume that \( \mathbf{S} \) is a supervisor for \( \mathbf{G} \) and that for simplicity \( X_\theta = X \).

Let \( \hat{\mathbf{S}} = (\Sigma, \hat{X}, \hat{\xi}, \hat{x}_0, \hat{X}_w) \) be a recognizer for \( L(\mathbf{S}, \mathbf{G}) \). Assume finally that \( \mathbf{S} \) is standard and weakly \( L(\mathbf{S}, \mathbf{G}) \)-reduced. Then we have

**Theorem 2 (Generalized quotient structure)**

There exists a control law
\[
\hat{\psi} : \Sigma \times \hat{X} \to \{ 0, 1, dc \}
\]
such that \( \hat{\mathbf{S}} = (\hat{\mathbf{S}}, \hat{\psi}) \) is a supervisor for \( \mathbf{G} \) with the property
\[
L(\mathbf{S}, \mathbf{G}) = L(\hat{\mathbf{S}}, \mathbf{G}) \quad L_{\text{red}}(\mathbf{S}, \mathbf{G}) = L_{\text{red}}(\hat{\mathbf{S}}, \mathbf{G})
\]
Furthermore, there is a cover \( \hat{\mathbf{C}} \) of \( \hat{\mathbf{S}} \) such that the corresponding reduced supervisor (based on \( \mathbf{C} \)) is isomorphic to \( \mathbf{S} \).

In applications the transition structure of a supervisor is typically much simpler than that of a recognizer for the closed-loop language that the supervisor synthesizes in conjunction with the DES that it controls. However, our result (like its prototype in Ramadge and Wonham 1983) captures the intuition that any 'reasonable' supervisor can be derived from (or is a high-level version of) a supervisor \( \mathbf{S} \) modelled on the grammatical structure of the supervisory task to be accomplished.

**Proof**

Theorem 2 will be proved in three steps. In (i) the first statement is proved, in (ii) a cover is constructed for \( \hat{\mathbf{S}} \) by use of the properties of \( \mathbf{S} \), and in (iii) it is shown that the corresponding reduced supervisor \( \hat{\mathbf{S}} \) is isomorphic to \( \mathbf{S} \). In this proof we write
\[
K := L(\mathbf{S}, \mathbf{G}) \quad L := L(\mathbf{G})
\]

(i) Let \( \hat{\mathbf{S}} = (\Sigma, \hat{X}, \hat{\xi}, \hat{x}_0, \hat{X}_w) \) be a recognizer for \( K \). We define a control function \( \hat{\psi} : \Sigma \times \hat{X} \to \{ 0, 1, dc \} \) as follows. Let \( \sigma \in \Sigma, x \in \hat{X} \). Then set
\[
\hat{\psi}(\sigma, x) = \begin{cases} 1 & \text{if } \sigma \in K \text{ and } \hat{\xi}(s, x_0) = x \& \sigma \in K \\ 0 & \text{if } \sigma \in K \text{ and } \hat{\xi}(s, x_0) = x \& \sigma \in L - K \\ dc & \text{otherwise} \end{cases}
\]
Since \( \hat{\mathbf{S}} \) is a recognizer for \( K \), two strings \( s, s' \in K \) such that \( \xi(s, x_0) = \xi(s', x_0) \) must be \( L \)-equivalent, namely \( \sigma \in K \text{ if and only if } \sigma \in K \). Thus the conditions for \( \hat{\psi} = 1 \) (respectively 0) are mutually exclusive, as they should be. Also, as \( K \) is controllable, case \( \hat{\psi} = 0 \) can arise only if \( \sigma \in \Sigma \). By these two remarks, \( \hat{\psi} \) is well defined. It is now evident that \( \hat{\mathbf{S}} = (\hat{\mathbf{S}}, \hat{\psi}) \) is a standard supervisor. Furthermore, one easily sees (by induction on length of strings) that
Finally

\[ L(S, G) = K = L(S, G) \]

and statement (i) of the theorem is proved.

(ii) We construct a cover for \( \hat{C} \) as follows. Let

\[ \hat{C} := \{ \hat{X}_x \mid x \in X \} \]

where

\[ \hat{X}_x := \{ \hat{z}(s, x_0) \mid s \in K \} \]

Since \( S \) is \( K \)-accessible, \( \hat{X}_x \neq \emptyset \); and since \( \hat{S} \) is \( K \)-accessible,

\[ \cup \{ \hat{X}_x \mid x \in X \} = \hat{X} \]

Since \( \hat{X}_x = \hat{X} \) the marking condition is satisfied. Next, fix \( x \in X \), \( s \in \Sigma \) and assume that for some \( \hat{z} \in \hat{X}_x \), \( \hat{z}(s, x) \neq \emptyset \). Now for some \( s \in K \),

\[ \hat{z} = \hat{z}(s, x), \quad x = \hat{z}(s, x_0) \]

Then \( \hat{z}(s, \hat{z}) = \hat{z}(s, x_0) \) and \( s \sigma \in K \). Thus \( \hat{z}(s, x_0) \), so that

\[ \hat{z}(s, x_0) = \hat{z}(s, x) = \hat{z}(s, x) \]

say. Therefore

\[ \hat{z}(s, \hat{z}) \in \{ \hat{z}(t, x_0) \mid t \in K \} = \hat{X}_x \]

That is, \( \hat{z}(s, \hat{z}) \in \hat{X}_x \), so we have proved the cover property

\[ (\forall s, \sigma) (\exists \hat{z})(\hat{z}(s, \hat{z}) \in \hat{X}_x) \]

It remains to check that \( \hat{C} \) is adapted to \( \hat{\psi} \). Fix \( z \in X \), \( s \in \Sigma \), \( \hat{z} \in \hat{X}_x \), and assume that \( \hat{\psi}(s, \hat{z}) \neq \hat{\psi}(s, \hat{z}) \). If \( \hat{\psi}(s, \hat{z}) = 1 \) then for some \( s \in K \) we have \( s \sigma \in K \) and \( \hat{z}(s, x_0) \neq \hat{z}(s, x_0) \). Since \( \hat{z} \in \hat{X}_x \), there is \( s' \in K \) such that \( \hat{z}(s', x_0) = z \) and \( \hat{z}(s', x_0), \hat{z}(s', x_0) = \hat{z}(s', x_0) \). Thus \( s' = s \mod K \) and so \( s \sigma \in K \) implies \( s' \sigma \in K \), hence \( \hat{\psi}(s, \hat{z}) = 1 \). Now for some \( t' \in K \), \( t'(s', x_0) = z \) and \( \hat{z}(t', x_0) = \hat{z}(s', x_0) \). Suppose \( \hat{\psi}(s, \hat{z}) = 0 \). Then there is \( t \in K \) with \( t \sigma \in L - K \) and \( \hat{z}(t, x_0) = \hat{z}(t, x_0) = z \). But now \( t \sigma \in L - K \) (together with the fact that \( K = L(S, G) \)) implies that \( \hat{\psi}(s, \hat{z}) = 0 \), a contradiction. We have shown that

\[ (\forall z, \hat{z} \in \hat{X}_x) (\forall s, \sigma \in \Sigma) (\exists \hat{z})(\hat{z}(s, \hat{z}) = 1 \land \hat{\psi}(s, \hat{z}) \neq \hat{\psi}(s, \hat{z}) \neq \hat{\psi}(s, \hat{z}) = 1) \]

from which it follows at once that

\[ (\forall z, \hat{z} \in \hat{X}_x) (\forall s, \sigma \in \Sigma) (\exists \hat{z})(\hat{z}(s, \hat{z}) \neq \hat{\psi}(s, \hat{z}) \land \hat{\psi}(s, \hat{z}) \neq \hat{\psi}(s, \hat{z}) \neq \hat{\psi}(s, \hat{z}) = 1) \]

as required.

(iii) It remains to construct a reduced supervisor \( S = (\Sigma, X, \hat{\psi}) \) based on \( \hat{C} \) that is isomorphic to \( S \). Let

\[ S = (\Sigma, X, \hat{\psi}, x_0, X) \]

Define \( X' := X, x_0' := x_0 \). Let \( x \in X \) and suppose \( \hat{z}(s, x) = y \) with \( \hat{z}(s, x_0) = x \) for some \( s \in K \). Since \( S \) is standard, such \( y \) exists. Then we know that \( \hat{z}(s, X_0) \subset X_0 \), and so we define \( \hat{z}(s, x) = y \). It only remains to set, for every \( s \in \Sigma \) and \( x \in X \), \( \hat{\psi}(s, x) = \psi(s, x) \), and the proof is complete.

\[ \square \]

In the foregoing proof the condition that the supervisor be weakly \( K \)-reduced is essential if the cover constructed for \( \hat{S} \) is to have the crucial feature of consistency with respect to control action.

2.4. Examples

2.4.1. Example 1

Let \( G \) have the state transition graph displayed below:

\[ \text{Figure 2.} \]

Here

\[ \Sigma = \{ \sigma, \beta, \gamma, \lambda \}, \quad \Sigma_c = \{ \lambda \} \]

\[ Q = \{0, 1, 2\}, \quad q_0 = 0, \quad Q_m = \{0\}. \]

To define \( S = (\Sigma, \psi) \) we take for \( S \) the same transition graph as for \( G \), but with the event 2 deleted, as shown below:

\[ \text{Figure 3.} \]

Thus \( X = \{0, 1, 2\}, x_0 = 0, X_m = X, \)

\[ \psi(\lambda, 0) = \delta c \]

\[ \psi(\lambda, 1) = 1 \]

\[ \psi(\lambda, 2) = 0 \]

As a cover for \( S \) we let \( C = \{ X_0, X_1 \} \), with

\[ X_0 = \{0, 2\}, \quad X_1 = \{0, 1\} \]
The corresponding reduced supervisor $S = (S, \psi)$ is given by the transition diagram and listing below:

![Diagram of S]

Figure 4.

\[ \dot{x}_0 = 0 \]
\[ \dot{y}(x, 0) = 0 \]
\[ \dot{y}(x, 1) = 0 \]

In this example there is no partition that reduces $S$, because the subsets of $X$ in the list \{0, 1\}, \{0, 2\}, \{1, 2\} are 'split' under the action of $x, y, \beta$ respectively; and the subset \{0, 1, 2\} is not consistent with respect to $\psi$. The quotient structure theorem of Ramadge and Wonham (1983) does not apply to $S$ because $S$ is not $K$-reduced: notice that $x = y^2 (\text{mod } K)$, but in $S$

\[ \zeta(x, 0) = 0 \neq 1 = \zeta(y^2, 0) \]

On the other hand, we claim that $S$ is weakly $K$-reduced; namely any string $s \in K$ such that

\[ \zeta(s, 0) = 2 \]  

(1)

(1) is true if $s \in s^* \beta^*$ for suitable $s'$; but for $t \in x \beta^*$,

\[ \zeta(t, 0) = \zeta(t, 1) = 0 \]

Thus $S$ exemplifies the generalized quotient structure of Theorem 2.

2.4.2. Example 2

Let $G$ have the state transition graph displayed below:

![Diagram of G]

Figure 5.

Here

\[ \Sigma = \{x, \beta, y, \lambda\}, \quad \Sigma_0 = \{\lambda\} \]
\[ Q = \{0, 1, 2, 3, 4\}, \quad q_0 = 0, \quad Q_m = \{0\} \]

To define the language $K = L(S, G)$ we adopt the transition graph of $G$ with the event 3, 4 deleted. A recognizer $S$ for $K$ is then obtained by merging states 2 and 3 to form a new state (2), with the result shown below:

![Diagram of S]

Figure 6.

A supervisor $S = (S, \psi)$ with $L(S, G) = K$ is provided by setting

\[ X = \{0, 1, (23), 4\}, \quad x_0 = 0, \quad X_m = X \]
\[ \psi(\lambda, (23)) = 0, \quad \psi(\lambda, 4) = 1 \]
\[ \psi(\lambda, 0) = \psi(\lambda, 1) = dc \]

We claim that an alternative supervisor $S = (S, \psi)$ with $L(S, G) = K$ is as displayed below:

![Diagram of S]

Figure 7.

\[ \dot{\psi}(\lambda, 0) = 1, \quad \dot{\psi}(\lambda, 1) = 0 \]

Indeed $S$ imposes exactly the required constraint that $\lambda$ be disabled following $y$, until the next occurrence of $x$. Clearly $S$ is of minimal state size. However, $S$ is not weakly $K$-reduced, since $\beta^2 = y (\text{mod } K)$ but

\[ \zeta(\beta^2, 0) = 0 \neq 1 - \zeta(y, 0) \]

By use of the algorithm of § 3 it was found that a minimal cover for $S$ contains at least 3 elements. That is, in general, reduction of a recognizer for $K$ by a cover need not always be effective in finding a reduced supervisor of minimal size. We shall return to this point in the next subsection.
2.5. Second generalized quotient structure theorem

For our second generalization of the quotient structure theorem of Ramadge and Wonham (1983) we shall drop the condition on the given supervisor $S$ that it be weakly $L(S,G)$-reduced, at the expense of increasing the complexity of the 'canonical' supervisor $\hat{S}$ from which $S$ is to be obtained by an appropriate cover.

As before, write $K = L(S,G), L = L(G)$. We construct a recognizer

$$\hat{S} = (\Sigma, \hat{X}, \hat{\zeta}, \hat{x}_0, \hat{X})$$

for $K$ as follows. First let

$$T = (\Sigma, Y, \eta, y_0, Y_m)$$

be a recognizer for the language

$$M := (L - K) \cup (K - (L - K))$$

By definition of a recognizer,

$$\bar{M} = \{ s \in \Sigma^* | \eta(s, y_0) \}$$

$$Y = \{ \eta(s, y_0) | s \in M \}$$

$$M = \{ s \in M | \eta(s, y_0) \in Y_m \}$$

Now set

$$Y_m = \{ \eta(s, y_0) | s \in L - K \}$$

$$\hat{X} = Y - Y_m$$

$$\hat{x}_0 = y_0$$

with

$$\hat{\zeta}(\sigma, \hat{x}) = \eta(\sigma, \hat{x})$$

iff $\hat{x} \in \hat{X}$ and $\eta(\sigma, x) \in \hat{X}$; otherwise $\hat{\zeta}$ is undefined. Notice that $1 \in K$ implies

$$\eta(1, y_0) = y_0 \notin Y_m$$

so that $\hat{x}_0 \notin \hat{X}$. We claim that

$$K = M - (L - K)$$

In fact, $K \subseteq L$ implies $M \subseteq L$, so that

$$M - (L - K) = \bar{M} \cap (L \cap K') = \bar{M} \cap K \subseteq K \cap K \subseteq K$$

while

$$K - (L - K) = K \cap (L - K) \subseteq K$$

and the claim is proved. Since $K$ is closed, it follows that $\hat{\zeta}(s, \hat{x}_0)$ iff $s \in K$, namely the automaton $\hat{S}$ is a recognizer for $K$. The special feature of $\hat{S}$ is that its states $\hat{x}$ represent the Nerode classes of $K$ 'split' relative to $L - K$, i.e. $s \in K$ and $\hat{\zeta}(s, \hat{x}_0) = \hat{\zeta}(t, \hat{x}_0)$ imply $s = t \mod (L - K)$. In particular, for this reason, it may be true that $|X| > |K|$. We now proceed as in § 2.3, but assume only that $S$ is standard. With $S, \hat{S}$ as just described, the statement of Theorem 2, and parts (i) and (iii) of its proof, may be repeated without change. In part (ii) of the proof it is necessary only to recheck that the cover $G$ is adapted to $\hat{S}$. Again, fix $x, \sigma \in \Sigma, \tilde{x}, \tilde{y} \in \hat{X}$, and assume that $\hat{\psi}(\sigma, \tilde{x}) \neq dc \neq \hat{\psi}(\sigma, \tilde{y})$. If $\hat{\psi}(\sigma, \hat{x}) = 1$ then it follows, as before, that $\hat{\psi}(\sigma, \hat{z}) = 1$. Now let $t' \in K$, $\hat{\zeta}(t', \hat{x}_0) = \hat{z}$ and $\hat{\zeta}(t', \hat{y}_0) = \hat{y}$. Suppose $\hat{\psi}(\sigma, \hat{y}) = 0$. Then there is $t \in K$ with $t \sigma \in L - K$ and $\hat{\zeta}(t, \hat{x}_0) = \hat{y}$. By the construction of $\hat{S}$, it follows that $t = t' \mod (L - K)$, and so $t \sigma \in L - K$ implies $t \sigma \in L - K$. As $\hat{\zeta}(t', \hat{x}_0) = \hat{z}$ and $K = L(S, G)$, we conclude that $\hat{\psi}(\sigma, \hat{z}) = 0$, a contradiction. With this our revised version of part (ii) of the proof of Theorem 2 is complete.

To illustrate the scope of this result we consider again Example 2 (§ 2.4.2). With a little calculation it can be verified that a recognizer $\hat{S}$ as described above can be taken as just the transition structure of $G$ with the transition $3 \rightarrow 4$ deleted; but this time we cannot merge states 2 and 3, as these states are not Nerode-equivalent mod $(L - K)$. To define $\hat{S}$ we assign the controls

$$\hat{\psi}(\sigma, 3) = 0, \quad \hat{\psi}(\sigma, 4) = 1$$

$$\hat{\psi}(\sigma, \hat{x}) = dc, \quad \hat{x} = 0, 1, 2$$

A cover $C = \{ \hat{X}_0, \hat{X}_1 \}$ is given by

$$\hat{X}_0 = \{ 0, 1, 2, 4 \}, \quad \hat{X}_1 = \{ 3 \}$$

The corresponding reduced supervisor is clearly isomorphic to the two-state supervisor $S$. We remark that, in general, a minimal-state supervisor $\hat{S}$ can always be taken to be standard, and will then satisfy the generalized quotient structure theorem of this section. By minimality, and the final observation of § 2.2, the cover $C = \{ X_i \}$ on which $S$ is based must have the 'non-redundancy' property

$$(\forall i, j) X_i \cap X_j \Rightarrow i = j$$

3. An algorithm for optimal reduction

In this section we assume that $L(G), L(S, G)$ are regular. From Theorem 1 we know that the synthesis of a minimal supervisor entails two steps: the search for a cover with a minimal number of elements, and the construction of a supervisor from that cover. We need the following definitions. Two states $x, y, x \in X$ are control-compatible if

$$\forall \sigma \in \Sigma \exists \psi(\sigma, x) \neq \psi(\sigma, y) \Rightarrow \hat{\psi}(\sigma, x) = \hat{\psi}(\sigma, y)$$

The (one-step) transition pairs of a state pair $(x, y) \in X \times X$ is the set of pairs $TP(x, y)$ defined by

$$TP(x, y) := \{ (x', y') | x', y' \in X, x' \neq y', (x, y) \in \Sigma \}$$

such that $\hat{\zeta}(\sigma, x') = \hat{\zeta}(\sigma, y')$.

Two states $x, y \in X$ are compatible if (i) $x, y$ are control-compatible, and (ii) $x \in X_m$ if $y \in X_m$ and (iii) all state pairs $(x', y') \in TP(x, y)$ satisfy (i) and (ii). A non-empty set of states $B \subseteq X$ is a block if every state pair $(x, y) \in B$ is compatible.

Compatible state pairs are determined by use of a merger table (cf. Paul and Unger 1959, Kohavi 1978). A table for five states is shown below:

```
  1 2 3 4 5
0 | 1 2 | 3 4
1 | 1 2
2 | 1 2 | 3 4
3 | 1 2
4 | 1 2
```

Figure 8.
Supervisor reduction in discrete-event systems

progress := true;
skip := true;
end;
if cell (s, t) \notin \text{boolean then}
okay := false;
if ((cell (s, t) = true) and (okay = true)
and (last = true)) then
begin
  cell (i, j) := true;
  progress := true
end
until ((last = true) or (skip = true))
end;

{set undetermined cells to true}
for a := 1 to l.c.d. do
  if cell (i, j) \notin \text{boolean then}
  cell (i, j) := true

end.

It is now the case that a state pair \((x_0, x_1)\) is compatible (i.e. a block) iff cell 
\((i, j) = true\). Blocks containing more than two states are constructed in a similar way.
All possible blocks must be listed. For an \(n\)-state supervisor there are at least about \(2^n\) blocks. The block list is searched for all singletons, pairs, triples, etc. of blocks until
a minimal cover is found. The computation is exponential in time and so is feasible
only for small values of \(n\).

From a given supervisor \(S = (S, \phi)\) with \(S = (\Sigma, X, \xi, x_0, X_a)\) we construct a
reduced supervisor \(\hat{S} = (\hat{S}, \hat{\phi})\) with \(\hat{S} = (\Sigma, Y, \eta, y_0, Y_a)\). Let \(C = \{B_0, B_1, ..., B_k\}\) be a
minimal cover with \(x_0 \in B_0\). The synthesis is carried out by the following procedure:
procedure synthesize; \{constructs a supervisor from a given cover\}
begin
  \{construct transition function \(\eta(\sigma, y)\}\)
  for \(i := 0 \text{ to } r \text{ do}\)
  \begin{align*}
    & \text{for each } \sigma \in \Sigma \text{ do} \begin{align*}
      & R := \{x' | x' = \xi(\sigma, x), x \in B_i\}
      \text{if } R \neq \emptyset \text{ then} \begin{align*}
        & \text{select } k \text{ such that } R \subset B_k;
        \eta(\sigma, y) := y_k
      \end{align*}
    \end{align*}
  \end{align*}

end;

\{construct control \(\hat{\phi}\)\)
for \(i := 0 \text{ to } r \text{ do}\)
\begin{align*}
  & \text{for each } \sigma \in \Sigma \text{ do} \begin{align*}
    & \text{if } x \in B_i \text{ such that } \phi(\sigma, x) \neq dc
  \end{align*}
\end{align*}

Let the states \(x \in X\) be labelled \(x_0, x_1, \ldots\). Each cell of the merger table corresponds to
a pair of distinct states \((x_i, x_j)\). For \(i > j\) let cell \((i, j)\) denote the cell in row \(i\), column \(j\) of
the table. The cells are filled in according to the following procedure.

procedure fillin; \{fills in merger table\}
begin
  for each pair \((x_i, x_j) \in X \times X\) with \(i > j\) do
  begin
    compatible := true;
    if \((x_i, x_j) \notin \text{not control-compatible})
    or \((x_i \in X_m\) and \(x_j \in X - X_m)\) or \((x_i \in X - X_m\) and \(x_j \in X_m))
    then compatible := false;
    if compatible = false then cell \((i, j)\) := false
    if compatible = true then
    begin
      form \(TP(x_i, x_j)\);
      \(TP(x_i, x_j) = \emptyset\) then \(cell (i, j) := true\)
      else \(cell (i, j) := TP(x_i, x_j)\)
    end
  end
end.

When the merger table is completed, cells containing transition pairs are assigned
an appropriate boolean value by the resolution procedure described in Paull and
Unger (1959) and Kohavi (1978). This process is termed \textit{resolving the merger table}.
Consider a cell \((i, j)\) that contains transition pairs \(TP(x_i, x_j)\). If any \((x', x) \in TP(x_i, x_j)\)
refers to a cell that contains the value \textit{false}, then cell \((i, j)\) is assigned the value \textit{false}. This
idea can be used to resolve the merger table. Let the cells containing transition pairs
be denoted by \{\cell \((i_1, j_1)\) \ldots \cell \((i_k, j_k)\)\}. The merger table is resolved by the following procedure:

procedure resolve; \{resolves the merger table\}
begin
  progress := true;
  while progress = true do
  begin
    progress := false;
    for \(a := 1 \text{ to } k \text{ do}\)
    begin
      okay := true;
      last := false;
      skip := false;
      if cell \((i_a, j_a)\) \notin \text{boolean then}
      repeat
        get transition pair \((x_i, x_i)\)
        from \(TP(x_i, x_i)\);
        if \((x_i, x_i)\) is the last such pair
        then last := true;
        if cell \((s, t)\) = false then
        begin
          cell \((i_a, j_a)\) := false;
        end
    end
  end
end.

\begin{align*}
  & \text{progress := true;}
  \text{skip := true;}
  \text{end;}
  \text{if cell (s, t) \notin \text{boolean then}}
  \text{okay := false;}
  \text{if ((cell (s, t) = true) and (okay = true)}
  \text{and (last = true)) then}
  \text{begin}
  \text{cell (i, j) := true;}
  \text{progress := true}
  \text{end}
  \text{until ((last = true) or (skip = true))}
\end{align*}

then \( \psi(\sigma, y) = \psi(\sigma, x) \)
else \( \psi(\sigma, y) = \Delta c \)
end.

A program based on these procedures has been tested on the VAX 11/780. For the 9-state supervisor of Example 1 in Ramadge and Wonham (1983) a reduction to 4 states was achieved. The computation generated 37 blocks and required 12 s of CPU time. For the 13-state supervisor of Lin (1984, p 34) a reduction to 6 states was achieved in 58 s of CPU time.

4. Conclusions

The main structural results provide a generalization of the supervisor reduction scheme and quotient structure theorem of Ramadge and Wonham (1983). Examples show that reduction by covers may indeed lead to supervisors with smaller state sets than those achievable by partitions. From a theoretical viewpoint, the generalized quotient structure theorems improve on the result of Ramadge and Wonham (1983) in placing less stringent requirements on the class of supervisors that are shown to be ‘quotients’. The fact that covers may improve on partitions depends essentially on the feature that the supervisor transition function and control law are generally partial functions; if both of these are total functions, any reduction by a cover can be shown to be equivalent to a reduction by a suitable partition with at most as many elements. In practice, of course, the exponential complexity of the computation of a minimal cover will rule out covers as a design approach in all but relatively simple cases. Nevertheless, in combination with other techniques of supervisor synthesis (notably decentralization and modularity) reduction by covers can play a useful role.

References


Start-up performance of different proportional–integral–anti-wind-up regulators

A. H. Glattfelder† and W. Schaufelberger‡

It has been found that proportional–integral–anti-wind-up regulators may produce unacceptable sluggish start-up in a large class of industrial plants, such as DC drives and temperature or pressure-control systems. In this paper the effect is clarified, better suited regulator structures are discussed in both continuous and discrete forms, and stability test are outlined.

1. The problem

In control-engineering practice many process models are of the ‘dominant first-order lag’ type, i.e.

\[
G_p(s) = \frac{1}{\prod_{k=1}^{m} (s - s_k)} \cdot \frac{1}{s - s_1}
\]

with

\[
s_k = -\sigma_k, \quad \sigma_k > 0 \\
s_k = -\sigma_k \pm j\omega_k, \quad \sigma_k >> \sigma_1 \text{ for } k = 2, \ldots, m
\]

Typical examples are speed-regulated DC motors, temperature-controlled stirred tank reactions or pressure-controlled compressed-air or gas systems.

Regulation is generally performed by a standard proportional–integral controller with anti-wind-up, as the control variable may saturate during start-up.

This eliminates overshoot, but may produce a very slow start-up, which is a surprise as the latter is not the case if \( \sigma_1 = 0 \), i.e. for an integrator plant with multiple small lags (cf. Glattfelder et al. 1983). The aim in this paper is to clarify this effect, to discuss better suited regulator structures and to outline how stability of such control loops may be tested.

2. Standard PI-ARW regulators

The continuous form (Fig. 1 (a)) (Hanus 1980) corresponds to one of the standard industrial regulator structures. The anti-reset-windup (ARW) feedback acts on the integrator input only. Its output \( u_i(t) \) is zero as long as

\[
u \leq u_i(t) \leq u^\uparrow
\]

and the regulator is linear. If \( u \) tends to be driven outside \( u \downarrow \) or \( u^\uparrow \) by \( e_i(t) \), an \( u_i(t) \neq 0 \) is generated by highgain \((k_i)\) feedback driving \( u_i \) near \( u_i \) or \( u^\uparrow \).

Received 13 December 1985.

† Corporate Research and Development, Sulzer Brothers Ltd, CH-8401 Winterthur, Switzerland.
‡ Department of Automatic Control and Industrial Electronics, Swiss Federal Institute of Technology, CH-8092 Zürich, Switzerland.