Factored Stochastic Trees:

A Tool for Solving Complex Temporal Medical Decision Models

GORDON B. HAZEN, PhD

The *stochastic tree* is a continuous-time version of a Markov-cycle tree, useful for constructing and solving medical decision models in which risks of mortality and morbidity may extend over time. Stochastic trees have advantages over Markov-cycle trees in graphic display and computational solution. Like the decision tree or Markov-cycle tree, stochastic tree models of complex medical decision problems can be too large for convenient graphic formulation and display. This paper introduces the notion of *factoring* a large stochastic tree into simpler components, each of which may be easily displayed. It also shows how the rollback solution procedure for unfactored stochastic trees may be conveniently adapted to solve factored trees. These concepts are illustrated using published examples from the medical literature.
Key words: stochastic trees; DEALE models; decision analysis; Markov-cycle trees; temporal medical decision modeling. **(Med Decis Making 1993;13:227–236)**

A *stochastic tree* may be regarded as a continuous-time version of a Markov-cycle tree,[1] or alternatively, as a multi-state DEALE model.[2] Stochastic trees were introduced by Hazen.[3] They are useful for constructing and solving medical decision models in which risks of mortality and morbidity may extend over time, possessing significant advantages over Markov-cycle trees in model display and solution. Like decision trees and Markov-cycle trees, stochastic trees may be rolled back to determine such quantities as mean lifetime or mean quality-adjusted duration.

Figure 1, adapted from Hazen[3] and based on Matchar and Pauker,[4] illustrates the typical ingredients in a stochastic tree*:

1. *Incremental impact states:* These are states such as Well, Post Big Stroke, and Post Small Stroke in figure 1, whose impacts are proportional to their durations. Each of these states has an associated *quality factor*, indicating life years earned for each year spent in the state. In figure 1, Well has quality factor 1.0, whereas Post Small Stroke has quality factor 0.80, and Post Big Stroke has quality factor 0.20. The implication is, for example, that each year spent in Post Big Stroke is worth only 0.20 year in state Well.

2. *Instantaneous impact states.* These are states such as Stroke, Big Stroke, and Small Stroke in figure 1, which have short durations (assumed zero for simplicity), but whose impacts can be felt through associated *tolls*. For example, in figure 1, Big Stroke has a toll of 4 months, meaning that passing through Big Stroke subtracts the equivalent of 4 months from total life years (due to the poor quality of life during a big stroke). Similarly, Small Stroke has an associated toll of 1 month. Stroke has no toll.

3. *Stochastic transition arrows.* These are the wavy arrows emanating from incremental impact states. These arrows are labeled with *rates*, indicating the average numbers of transitions per unit time along them. The higher the rate, the more likely it is that the indicated transition will occur first. For example, in figure 1, the patient remains in state Well until either a stroke occurs or death occurs from another cause. Stroke occurs at rate $\mu_{stroke} = 0.05$ stroke per year, whereas death occurs at rate $\mu_o + \mu_{excess} = 0.01106 + 0.065 = 0.07606$ occurrence per year (data from Matchar and Pauker[4]).

4. *Probabilistic transition arrows.* These emanate from instantaneous impact states, are labeled with probabilities, and are identical in function to arrows from chance nodes in a decision tree.

5. *Repeated states.* These are the states whose borders are dashed instead of solid. Transition to one of these indicates that the previous state with the same name is to be revisited. For example, in figure 1, the patient remains in the state Post Small Stroke until death occurs or until transition back to the previous state Stroke occurs. We call stochastic trees with repeated states *cyclic*, and those without repeated states *acyclic*.

When quality factors and transition rates are given, mean quality-adjusted duration beginning at a particular state may be calculated from mean quality-ad-

Received January 16, 1992, from the Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois 60208-3119. Revision accepted for publication October 23, 1992. Supported in part by grant SES-9112854 from the National Science Foundation.

Address correspondence and reprint requests to Dr. Hazen.

*Graphic conventions in this and subsequent diagrams differ slightly from those in Hazen's earlier report.

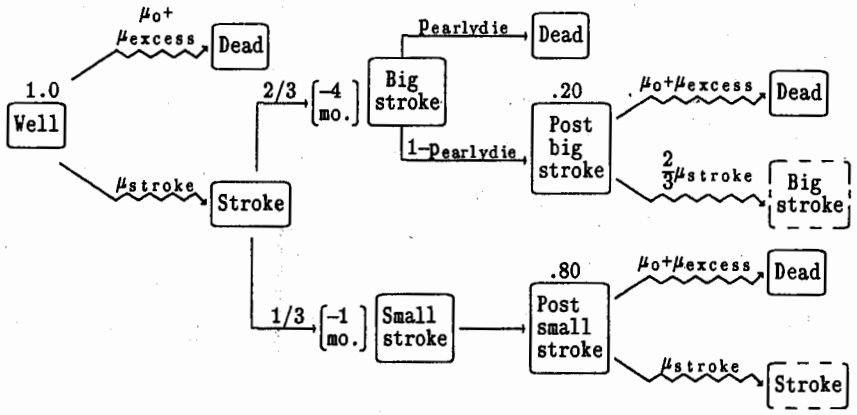


FIGURE 1. A stochastic tree model of recurrent stroke.

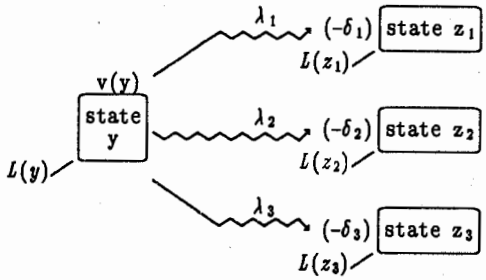


FIGURE 2. Rollback at an incremental impact state.

justed durations beginning at immediately subsequent states. The procedure at incremental impact states is as follows. Suppose that from state y, having quality factor v(y), transition at rate λ_i may occur to any one of the subsequent states z_i, with associated toll δ_i. Figure 2 depicts this situation when there are three possible subsequent states. Given the mean quality-adjusted durations $L(z_i)$ associated with subsequent states z_i, the corresponding quantity L(y) can be calculated using the formula

$$L(y) = \frac{v(y) + \Sigma_i(-\delta_i + L(z_i))\lambda_i}{\Sigma_i\lambda_i} \quad (1)$$

(For a derivation, see Hazen.[3]) For instantaneous impact states, the same probabilistic averaging procedure used in decision trees is appropriate. In these ways, mean quality-adjusted durations may be calculated for all states in an acyclic stochastic tree, beginning at the terminal states and proceeding recursively backwards to the initial state in the tree. In other words, acyclic stochastic trees may be *rolled back* to determine mean quality-adjusted duration.

Mean quality-adjusted durations can be obtained in cyclic stochastic trees by repeatedly rolling back the tree until the calculated L(y) values stabilize. This procedure is known as *value iteration*, or the *method of successive approximations*. Examples of value iteration are given below and by Hazen.[3]

Figure 3 illustrates this rollback procedure for an acyclic stochastic tree taken from Hazen, and based originally on the cancer/AIDS model of Roach et al.[5] The italicized quantities attached to each state are the L(y) values in equation 1 computed using $\mu_o = 0.014191$, $\lambda_c = 0.0325$, $\lambda_a = 0.10$, $\mu_c = 0.3081$, $\mu_a = 0.9979$ occurrences per year. For example, beginning in the No Disease state, mean subsequent quality-adjusted duration is 7.475 years, whereas it is only 0.4857 years beginning in the AIDS state.

The purpose of this report is to present a procedure, called *stochastic factoring*, for the formulation and display of large stochastic trees. Factored stochastic trees are split into components, which we call factors, each of which is itself a simpler stochastic tree. Formulation and display of the simpler factor trees is usually straightforward. Many complex processes may be thought of as several simpler subprocesses unfolding

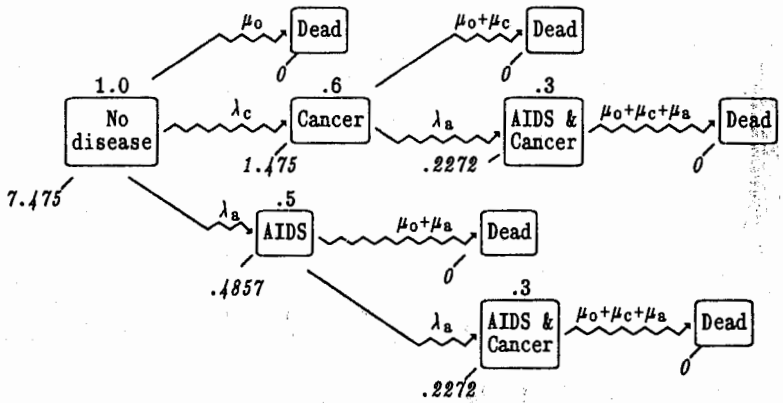


FIGURE 3. A stochastic tree, illustrating rollback.

in parallel. Often these parallel subprocesses may be thought of as factors in a large stochastic tree.

We also show how the rollback procedure for unfactored stochastic trees may be conveniently adapted to solve factored stochastic trees. Illustrations are given using published examples from the medical literature. Finally, comparison is made with techniques for displaying large Markov-cycle trees, and advantages of factored stochastic trees in model formulation and presentation are pointed out.

Factored Displays of Stochastic Trees

We introduce the notion of factoring stochastic trees by reconsidering the Cancer/AIDS example of figure 3. Notice that there are three sources of mortality in this model: cancer, AIDS, and other sources, with corresponding mortality rates μ_c, μ_a, μ_o. In fact, the model contains three independent subprocesses, unfolding in parallel: cancer, AIDS, and background mortality. These subprocesses may themselves be depicted as stochastic trees, and we do so in figure 4. Notice that the state of the overall process is actually determined by the states of the three subprocesses. For example, the triple (No Disease, No Disease, Alive) of subprocess states corresponds in the original tree to the state No Disease, and the triple (No Disease, AIDS, Alive) corresponds to the AIDS state. Moreover, the triple (Cancer, Dead, Alive) (or any other triple with one or more Dead components) corresponds to the state Dead in the original tree.

The quality factors in the original tree may also be obtained from the subprocess quality factors if we follow the approach of Roach et al.[5] and postulate that the overall quality factor is the product of the corresponding subprocess quality factors:

$$v(y_1, y_2, y_3) = v_1(y_1)v_2(y_2)v_3(y_3)$$

For example, v(Cancer + AIDS) = v(Cancer, AIDS, Alive) = (0.6)(0.5)(1) = 0.3, as desired; and the quality factor assigned to any triple with one or more Dead components is zero.†

Whenever we can depict a stochastic tree as a collection of subtrees unfolding in parallel such that the state in the overall tree can be recovered from the states of the subtrees, then we say we have *factored* the original tree. We call the subtrees *factors* of the original tree, and say that the original tree is the *product* of its factors.

There are many cases in which the number of health

†This method of assigning quality factors to product states is merely illustrative. In general, any coherent assignment of quality factors to product states is permissible, so long as product states with "Dead" components are assigned quality factor zero. Quality factors may even be assigned directly to product states, without attaching quality factors to states in the factors.

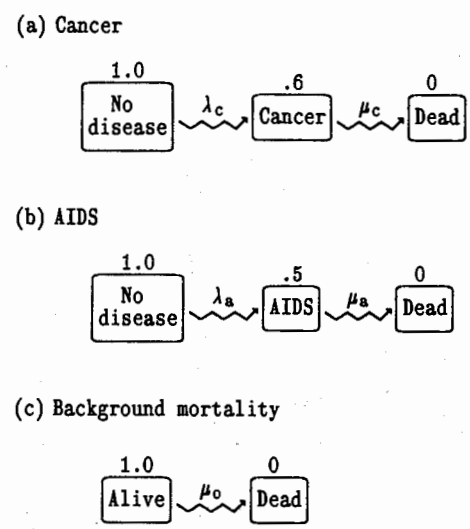


FIGURE 4. Factored representation of competing cancer and AIDS risks.

states in a stochastic-process model of medical management becomes large enough to impede formulation or display of the model. In fact, the usual reason is that a state descriptor is required that indicates the simultaneous status of several parallel subprocesses. A case in point is the discrete-time Markov-chain model used by Hillner et al.[6] to help analyze the use of postmenopausal estrogens in the prevention of osteoporosis. Hillner et al. formulate a Markov chain with the following states (from their table I):

Well
Uncomplicated fracture
Postfracture well
Uncomplicated hysterectomy
Posthysterectomy well
Endometrial cancer
Disabled
Nursing home
Posthysterectomy and uncomplicated fracture
Posthysterectomy and postfracture
Uncomplicated hysterectomy and postfracture
Disabled and endometrial cancer
Disabled and uncomplicated hysterectomy
Disabled and posthysterectomy
Disabled and new uncomplicated fracture
Endometrial cancer and uncomplicated fracture
Endometrial cancer and postfracture
Dead

These authors do not graphically portray the process. However, it may be seen that this 18-state formulation is really an attempt to desribe the simultaneous unfolding of three parallel subprocesses: fracture and its consequences; endometrial cancer and hysterectomy; and death due to natural causes. We supply a stochastic-tree description of these subprocesses in figure 5, based on our interpretation of Hillner et al.'s formulation.

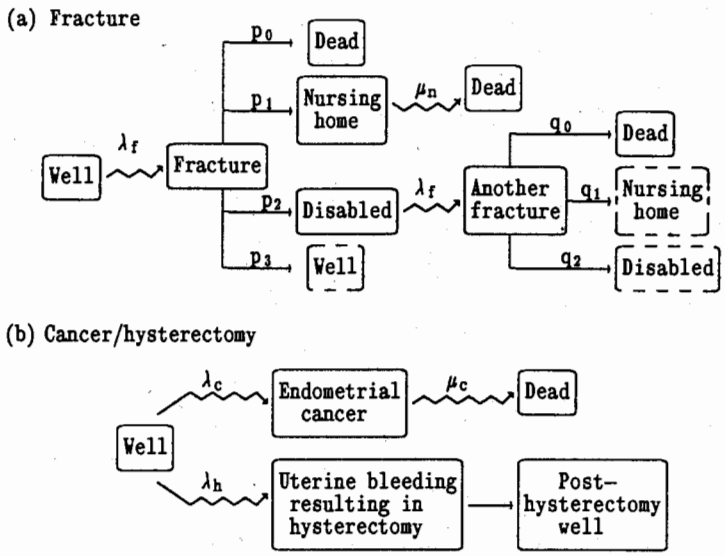


FIGURE 5. Factored stochastic tree representation of the use of postmenopausal estrogens in prevention of osteoporosis.

Notice that the Fracture factor is a cyclic stochastic tree: from the state Fracture, one may return to the previous state Well, as is indicated by the repeated state. Also, from Another Fracture one may revisit Disabled, or transit to the previous state Nursing Home. (The reason for the state Another Fracture is not to allow repeated fractures, which is possible already by cycling from Fracture to Well and back again to Fracture, but to allow fractures while preventing return to the Well state once Disabled has been visited.)

We have not specified the values of probabilities and transition rates in figure 5. Suffice it to note that postmenopausal estrogen decreases fracture rates and increases the rate of endometrial cancer and hysterectomy. The state of the product process is a triple (x,y,z) indicating the individual states x,y,z of the three factor processes. Any triple having "Dead" as one of its components is assumed to have quality factor 0. Thus, although there are $6 \times 5 \times 2 = 60$ product states, only $5 \times 5 \times 1 = 20$ of them have nonzero quality factor. The model is therefore comparable in size to the original Hillner et al. Markov chain. For computational purposes, further reduction in the number of states can be accomplished, as we shall see below.

Factoring stochastic trees in this manner has clear advantages for model presentation. Its advantages for model formulation are more subtle but nevertheless real. Considering one factor at a time reduces complexity, and thereby reduces the chance of modeling errors. In figure 5, one may, for example, focus exclusively on the modeling of fracture and its consequences, without immediately having to worry about how to incorporate Cancer/Hysterectomy transitions, which are irrelevant to fracture status. Similarly, one may formulate the Cancer/Hysterectomy factor without paying attention to Fracture transitions.

Sometimes, however, there are cross-factor dependencies that must be accounted for. For example, none of the original Hillner et al. states given above allows for the possibility of either hysterectomy or endometrial cancer while in a nursing home. This can be modeled in figure 5 by letting the rates λ_c and λ_h in the Cancer/Hysterectomy factor depend on the state of the Fracture factor. The rates λ_c, λ_h would be zero when the Fracture state was Nursing Home, but otherwise would take on their usual values.

A model that requires more complex cross-factor dependencies is given in figure 6, which is based on the embolism/hemorrhage model of warfarin efficacy for dilated cardiomyopathy constructed by Tsevat et al.[7] The anticoagulant drug warfarin reduces the rates of systemic and pulmonary embolism, but increases the rate of systemic hemorrhage. Does the benefit of reducing the embolism rate offset the increased rate of hemorrhage? Quality-adjusted lifetime calculations based on figure 6 can be used to address this question.

The model contains a number of cross-factor dependencies. For example, the rate λ_{se} of systemic embolism depends on whether anticoagulant status (AS) is Warfarin or not. The rates λ_{pe} of pulmonary embolism and λ_{sh} of systemic hemorrhage also depend on anticoagulant stutus. Another cross-factor dependency concerns warfarin administration strategy. If the patient is not taking warfarin and an embolism occurs, the strategy is to immediately place the patient on warfarin. This is indicated in the systemic embolism factor by the reference mark (+) at the *head* of the arrow leading to embolism. A reference mark such as (+), or some other parenthesized character at the head of a transition arrow, indicates that this transition may cause another transition in some other factor. The transitions caused are indicated by the presence of the same character at the *tails* of transition arrows somewhere else in the diagram. In this case, there is a (+) at the tail of the arrow leading from "No Warfarin" to "Warfarin" in the anticoagulant status factor. (The arrow is dashed because the transition cannot otherwise occur.) Therefore, the transition from No Warfarin to Warfarin occurs as soon as the systemic embolism factor enters the Embolism state. As can be seen, the same transition would also be forced whenever the pulmonary embolism factor enters the Embolism state. On the other hand, a systemic hemorrhage while taking warfarin causes temporary discontinuance of the drug, as indicated by the reference mark (−).

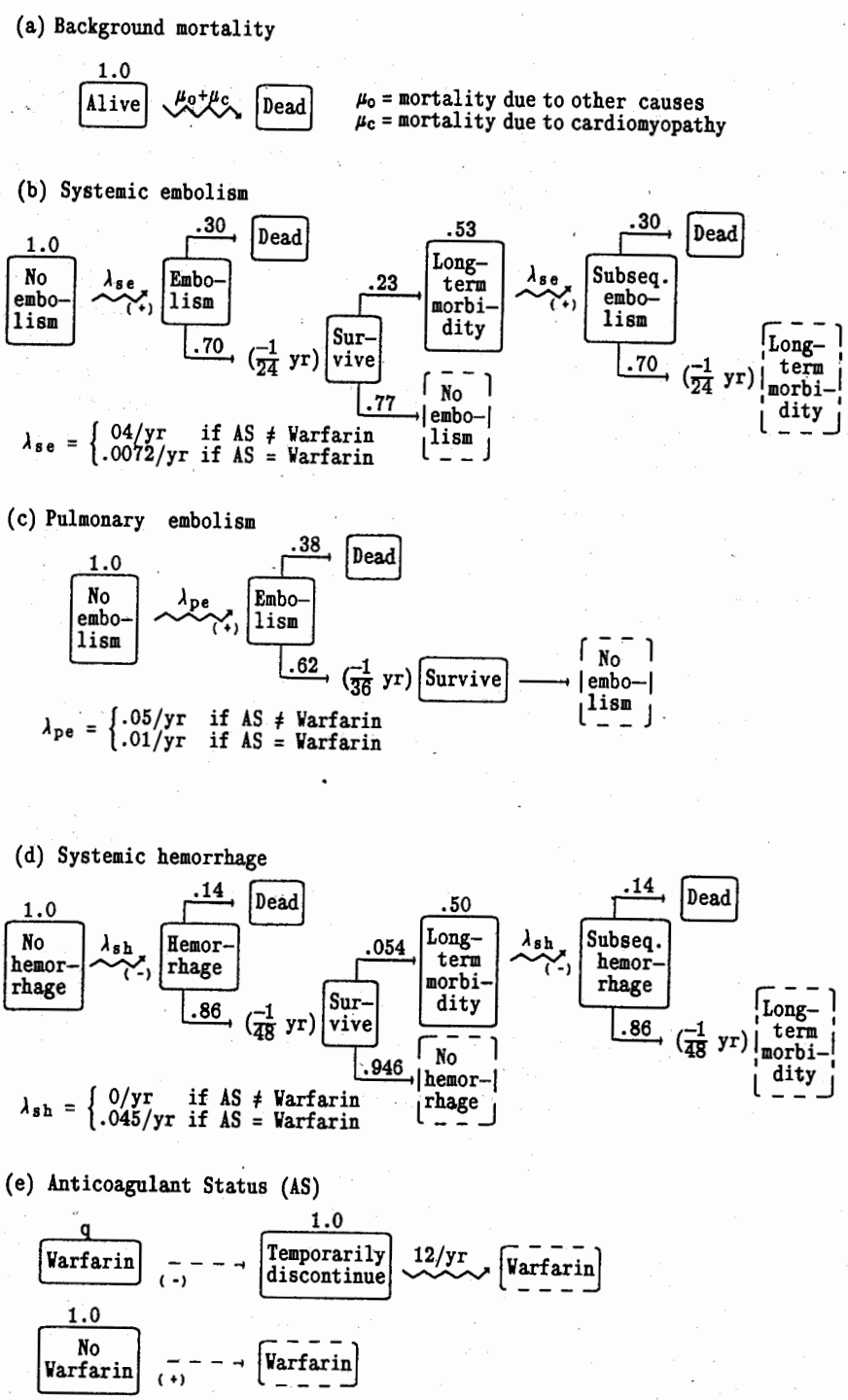


FIGURE 6A. Factored stochastic tree model for embolism/hemorrhage during treatment with Warfarin (factors a, b and c).

FIGURE 6B. Factored stochastic tree model for embolism/hemorrhage during treatment with Warfarin (factors d and e).

Rolling Back Factored Stochastic Trees

As the previous section illustrates, stochastic trees that are complex enough to require a factored display will usually have so many states that the graphic rollback procedure mentioned above becomes unwieldy to execute by hand: For example, the five-factor embolism/hemorrage model just discussed has $1 \times 5 \times 3 \times 5 \times 3 = 225$ nonfatal product states. For such large models, one would typically encode and execute the rollback algorithm in come convenient programming language. The purpose of this section is to show how the rollback solution procedure for unfactored stochastic trees can be conveniently adapted to solve factored stochastic trees. We also suggest a state-elimination strategy to speed up the rollback procedure.

STATE ELIMINATION

First, the state-elimination strategy: It is often pos-

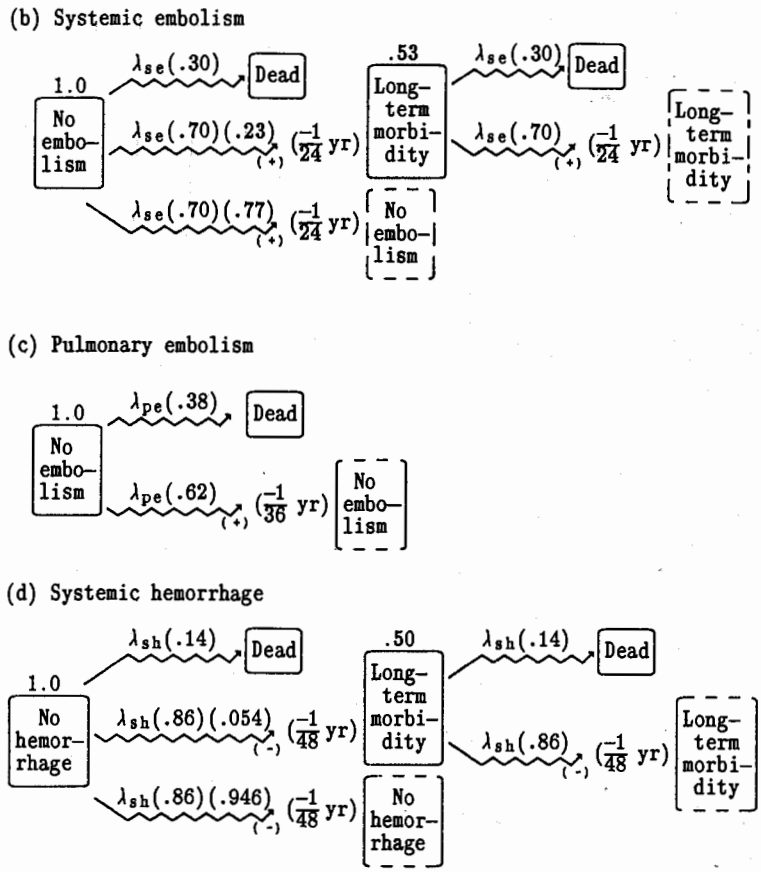
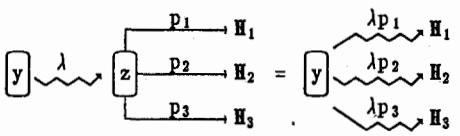


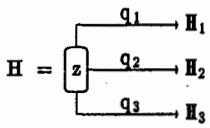
FIGURE 7. The results of state elimination in the embolism/hemorrhage model.

sible to eliminate instantaneous states from each factor of a stochastic tree using the decomposition rule given by Hazen.[3] In graphic form the rule is



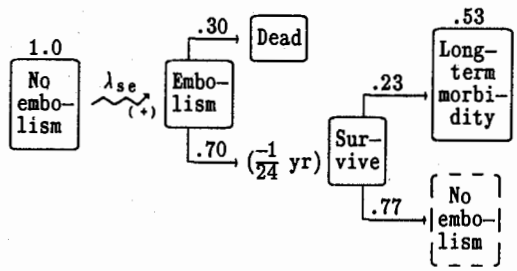
(2)

which has the following meaning. Suppose instantaneous state z can be entered from state y at rate λ, and from z one transits to subtree $\mathbf{H}_i$ with probability p_i. Then z may be eliminated and transitions made from y directly to subtree $\mathbf{H}_i$ at rate λp_i. The rule holds even if there are other possible transitions out of y. Using rule 2 and standard methods for combining chance nodes, one can eliminate from a stochastic tree all instantaneous states except initial ones not preceded by any stochastic transitions. The result would be a stochastic tree of the form

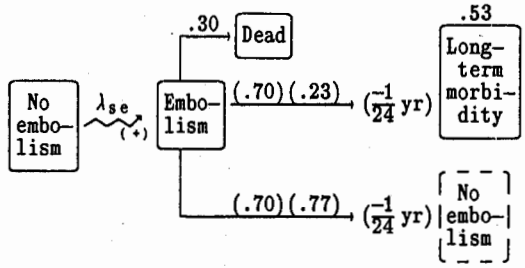


where z is an initial instantaneous state and the subtrees $\mathbf{H}_i$ contain no instantaneous states. The mean quality-adjusted duration L(**H**) associated with **H** is then $\Sigma_i q_i L(\mathbf{H}_i)$, so L(**H**) is determined once the subtrees $\mathbf{H}_i$ are rolled back to obtain $L(\mathbf{H}_i)$.

The motivation for state elimination is that the value-iteration procedure requires work proportional to the number of nonfatal product states. State elimination can therefore significantly speed rollback. Figure 7 illustrates how state elimination may be done in three of the five factors of the embolism/hemorrhage model. Consider, for example, the two instantaneous states Embolism and Survive in the Systemic Embolism factor of figure 6. Considering only the portion



of the Systemic Embolism factor, one may first combine the two chance nodes by multiplying probabilities in the standard way to obtain



making sure to transfer the −1/24-year toll to the subsequent states. Next one may apply the decomposition rule 2 to eliminate the Embolism state, and figure 7B results. Notice that the triggering symbol (+) is inherited by each stochastic arrow created.

The remaining state eliminations are performed in the same fashion. After state elimination, there are only $1 \times 2 \times 1 \times 2 \times 3 = 12$ nonfatal product states, compared with the initial 225 product states of figure 5, quite a dramatic reduction.

ROLLBACK

Next we turn to the value-iteration procedure for rollback. The procedure has a relatively intuitive interpretation in the factored tree setting. Assume that all chance nodes have been eliminated as just described, and suppose that for product state y, $\lambda_{ij}(y)$ is the transition rate from state y associated with the jth transition in the ith factor. Moreover, let $\varphi_{ij}(y)$ be the new

product state to which this transition leads, and let $\delta_{ij}(y)$ be the toll, if any, associated with this transition. For example, in the embolism/hemorrhage model of figures 6 and 7, number the factors in order from i = 0 to i = 4, and abbreviate the names of the factor levels ("No embolism," "Long-term morbidity," and so on) using the initial letters (N, L, and so on, respectively). Consider the product state

$$y = \text{ANNNN}$$

which abbreviates (Alive, No embolism, No embolism, No hemorrhage, No warfarin). Then

$$\lambda_{01}(\text{ANNNN}) = \mu_o + \mu_c$$

and

$$\lambda_{11}(\text{ANNNN}) = \lambda_{se}(0.30) = (0.04)(0.30)$$

$$\lambda_{12}(\text{ANNNN}) = \lambda_{se}(0.70)(0.23) = (0.04)(0.70)(0.23)$$

$$\lambda_{13}(\text{ANNNN}) = \lambda_{se}(0.70)(0.77) = (0.04)(0.70)(0.77)$$

and so on for the other factors. Moreover,

$$\varphi_{01}(\text{ANNNN}) = \text{DNNNN (fatal)}$$

and

$$\varphi_{11}(\text{ANNNN}) = \text{ADNNN (fatal)}$$

$$\varphi_{12}(\text{ANNNN}) = \text{ALNNW}$$

$$\varphi_{13}(\text{ANNNN}) = \text{ANNNW}$$

Notice that transitions 2 and 3 in factor 1 result in the change N → W "No warfarin" to "Warfarin" in factor 4. The tolls associated with these transitions are

$$\delta_{01}(\text{ANNNN}) = 0$$

and

$$\delta_{11}(\text{ANNNN}) = 0$$

$$\delta_{12}(\text{ANNNN}) = 1/24 \text{ yr}$$

$$\delta_{13}(\text{ANNNN}) = 1/24 \text{ yr}$$

Using this notation, value iteration for factored stochastic-tree rollback can be formulated as follows:

Initialize: For each product state y:
  $L_0(y) \leftarrow 0$ (Set initial quality-adjusted durations to zero)
  For each factor i:

$$w_i(y) \leftarrow \frac{\Sigma_j \lambda_{ij}(y)}{\Sigma_i \Sigma_j \lambda_{ij}(y)} \tag{3}$$

Repeat for n = 1,2,3, . . . until the values $L_n(y)$ stabilize:
  For each product state y:
    For each factor i:

$$L_{in}(y) \leftarrow \frac{w_i(y)v(y) + \Sigma_j \lambda_{ij}(y)(-\delta_{ij}(y) + L_{n-1}(\varphi_{ij}(y)))}{\Sigma_j \lambda_{ij}(y)} \tag{4}$$

    $L_n(y) \leftarrow \Sigma_i w_i(y) L_{in}(y)$

This algorithm is a straightforward extension of our original value-iteration algorithm. Notice in formula 3 that $w_i(y)$ is the proportion of total transition rate out of product state y that occurs in factor i. The assignment 4 is just our original rollback formula 1 [with $w_i(y)v(y)$ replacing $v(y)$] performed within factor i. One such calculation is done for each factor, and the results are weighted using the $w_i(y)$ to obtain the quality-adjusted duration $L_n(y)$ associated with y. This calculation is done for each product state y. The whole process is then repeated until convergence occurs.

We performed value iteration on the embolism/hemorrhage model of figures 6 and 7. For the background mortality rate μ_o, we used the value 0.07685/year, approximately correct for a 75-year-old white male. We had to estimate the excess mortality μ_c due to cardiomyopathy, since Tsevat et al. did not specify the value they used. We found that μ_c = 0.20/year gave quality-adjusted durations very similar to the values given by Tsevat et al. Moreover, to be consistent with Tsevat et al., the quality factor v(y) assigned to any product state y must be the minimum of the quality

Table 1 • Value Iteration for the Embolism/Hemorrhage Model of Figures 6 and 7

State	Cycle 0	1	2	4	6	10
ANNNN	0	2,720	3.170	3.234	3.243	3.245
ANNNW	0	2.946	3.048	3.391	3.430	3.435
ANNNT	0	0.081	2.940	3.363	3.421	3.429
ANNLN	0	1.357	1.594	1.629	1.634	1.635
ANNLW	0	1.471	1.525	1.701	1.722	1.725
ANNLT	0	0.040	1.468	1.687	1.717	1.722
ALNNN	0	1.439	1.690	1.727	1.732	1.733
ALNNW	0	1.560	1.616	1.803	1.825	1.828
ALNNT	0	0.043	1.556	1.788	1.820	1.825
ALNLN	0	1.357	1.594	1.629	1.634	1.635
ALNLW	0	1.471	1.525	1.701	1.722	1.725
ALNLT	0	0.040	1.468	1.687	1.717	1.722

factors $v_i(y_i)$ associated with component states y_i.

The numerical results of value iteration are presented in table 1. Convergence to three decimal places was achieved after ten cycles. The resulting values are mean quality-adjusted durations beginning in the indicated product state. The two values of interest are the 3.245 years associated with the initial no-warfarin state ANNNN (alive, no embolism or hemorrhage, no warfarin) and the 3.435 years associated with the initial warfarin state ANNNW (alive, no embolism or hemorrhage, warfarin). The difference 3.435 − 3.245 = 0.190 years = 69 days is the mean quality-adjusted benefit of warfarin.

Because the value-iteration process converges but does not terminate, the quality-adjusted durations at cycle 10 are only approximately correct. However, one can develop bounds on the maximum possible error after each cycle. We discuss this in the appendix. The cycle-10 values in table 1 are accurate to within 1.729×10^{-4}.

ROLLBACK USING AN ERLANG MORTALITY FACTOR

All the stochastic trees described above used transition rates that were stationary, so, for example, age-dependent mortality rates were not allowed. Using constant instead of the more accurate age-dependent rates can significantly affect quality-adjusted lifetime calculations. When accurate age-dependent mortality approximations are desired in a factored stochastic tree model, it is quite natural to use the Erlang mortality model introduced by Hazen[3] as a background mortality factor. We illustrate this procedure using the embolism/hemorrhage model of figures 6 and 7. The background mortality factor has mortality rate $\mu_o + \mu_c$, where μ_c is mortality rate due to cardiomyopathy and μ_o is mortality rate due to other causes. We wish to make the latter age-dependent. To this end, we split off cardiomyopathy mortality into a sixth factor of its own, and replace the background mortality factor by an Erlang factor with the appropriate rates and number of stages. Figure 8 depicts the result.

To compare with the data of Tsevat et al.,[7] we wish to calculate mean quality-adjusted durations for 35-, 55-, and 75-year-old white males. From Hazen,[3] we see that background mortality would require n = 7, n = 3, and n = 1 stages, respectively, with corresponding stage transition rates $\mu_o = 0.16985$, $\mu_o = 0.1300$, and $\mu_o = 0.07685$. Mean quality-adjusted durations were calculated by value iteration, and the results are shown in table 2. There, "Warfarin" indicates the policy of starting a patient on warfarin treatment and temporarily discontinuing if there is a hemorrhage. Similarly, "No Warfarin" indicates a policy of starting with no warfarin, but beginning warfarin therapy if there is an embolism. The quality factor q associated with warfarin use was taken to be 1.0 (no decrement in quality of life due to warfarin use).

Also shown in table 2 are quality-adjusted durations when only one type of embolism, either systemic or pulmonary, is allowed. These latter calculations are easily performed by simply omitting the other embolism factor (or equivalently, setting all its rates to zero). The same calculations were performed by Tsevat et al.,[7] and the results were qualitatively similar.

The number of nonfatal product states for the 35-

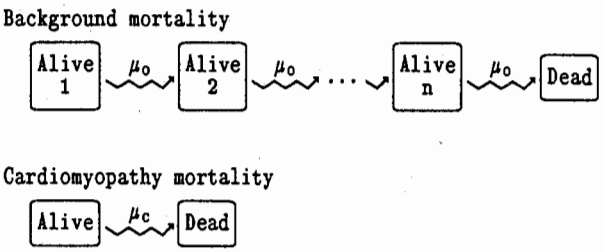


FIGURE 8. Revised mortality factors for the embolism/hemorrhage model.

Table 2 • Quality-adjusted Durations for the Embolism/Hemorrhage Model

	Warfarin (Years)	No Warfarin (Years)	Net Gain Warfarin (Days)
Age 35 years			
Combined	4.654	4.330	118
Systemic embolism only	4.739	4.641	36
Pulmonary embolism only	4.714	4.569	53
Age 55 years			
Combined	4.417	4.130	105
Systemic embolism only	4.490	4.405	31
Pulmonary embolism only	4.469	4.342	46
Age 75 years			
Combined	3.435	3.245	69
Systematic embolism only	3.482	3.427	20
Pulmonary embolism only	3.469	3.386	30

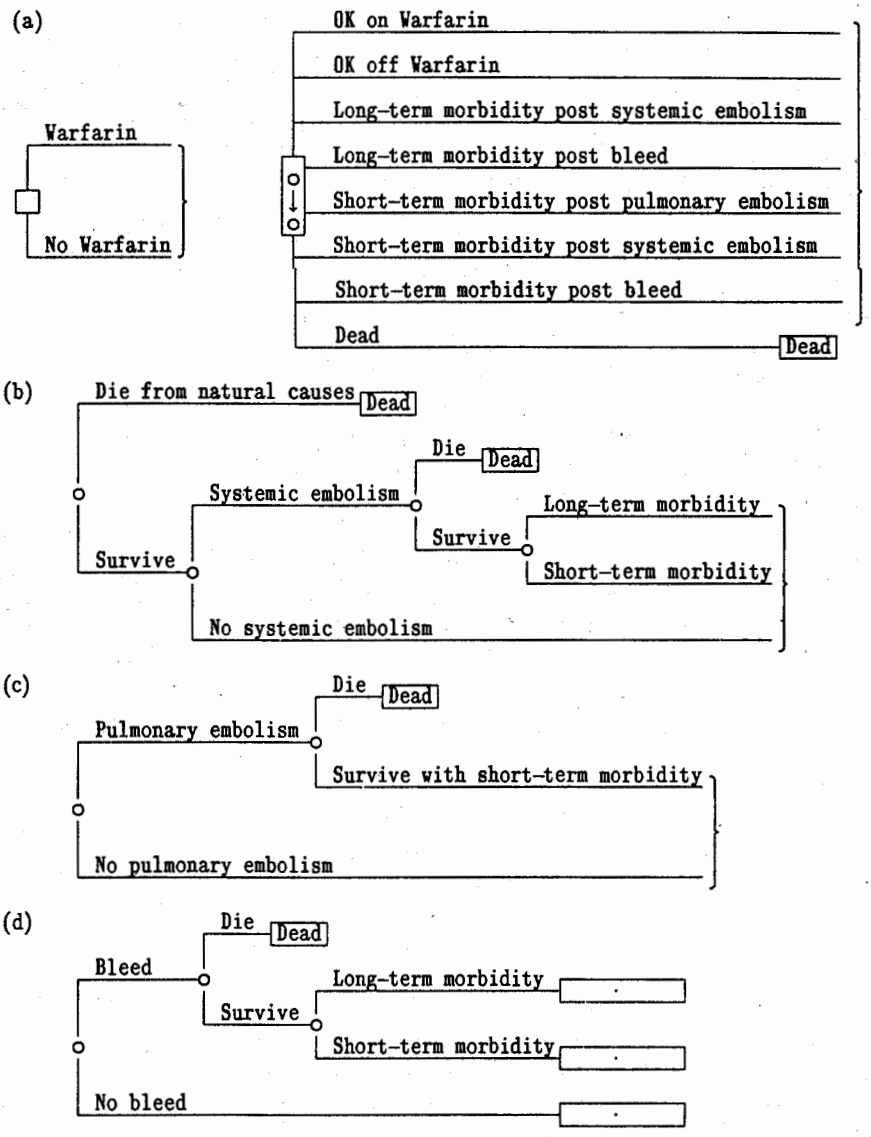


FIGURE 9. Markov-cycle tree model of embolism/hemorrhage.

year-old calculations was $7 \times 2 \times 1 \times 2 \times 3 \times 1 = 84$, and value iteration required 15 cycles to terminate. For the 55-year-old calculations, there were $3 \times 2 \times 1 \times 2 \times 3 \times 1 = 36$ product states and 12 cycles were required. The 75-year-old calculations involved only 12 product states and ten cycles, and were identical to those presented in table 1.

Comparison with Markov-cycle Trees

How does the factored stochastic-tree approach compare with previous techniques that have been used to display large Markov-cycle trees? The comparison is difficult to make because there are no systematized display procedures for large Markov-cycle trees. It may be instructive to compare the original cycle tree diagram for the embolism/hemorrhage model of Tsevat et al.[7] with its stochastic tree representation in figure 6. Tsevat's Markov-cycle tree diagram is reproduced in figure 9. Its interpretation is as follows. First, the drug strategy (Warfarin versus No Warfarin) is decided. Then, beginning in one of the "OK" states, the patient transits successively through parts (b), (c), (d) of the tree, afterwards returning to part (a) of the diagram to repeat the cycle (which is one month long). The factor structure for this model can be clearly seen in figure 9: part (b) includes what we have called the background mortality factor and the systemic embolism factor; part (c) is the pulmonary embolism factor; part (d) is the systemic hemorrhage factor.

Like the factored stochastic-tree display, the graphic feasibility of this cycle-tree diagram hinges on the fact that the same event possibilities occur in parts (b), (c), (d) regardless of the paths taken through the preceding parts. Unlike the stochastic-tree display, all transition probabilities, quality factors, tolls, and triggers have been left out of the cycle tree, resulting in a less cluttered appearance. However, should it be desired, the same omissions would also reduce clutter in a stochastic-tree display (e.g. figure 5).

There are structural elements of the stochastic-tree

model that apparently cannot be represented in the cycle-tree diagram. For example, the cycle tree does not specify which of the branches in (a) should follow particular transition paths in (b), (c), (d). If there is an embolism event with long-term morbidity in (b), but no pulmonary embolism in (c) and no bleeding in (d), then presumably transition is made to "Long-term morbidity post systemic embolism," although the diagram by itself does not indicate. But what should happen if *both* systemic and pulmonary embolism occur in (b), (c)?

These specific comparisons may not be fair, because there might well be other better methods of depicting factor structure in Markov-cycle trees. The point, however, is that there is currently no standard method available. Although it might be possible to adapt some of the conventions for displaying factored stochastic trees, the relatively more cumbersome graphic nature of the Markov-cycle tree compared with the stochastic tree[3] would, in my opinion, be an obstacle to such an effort.

Conclusion

This paper has introduced the notion of stochastic factoring, a graphic technique for the formulation and display of large stochastic trees. It has also illustrated how the stochastic-tree solution method of value iteration may be conveniently carried out in the factored setting. The factored stochastic tree is a complete but parsimonious tool for specifying the probabilistic structure of a stochastic medical decision model. By allowing model assumptions to be presented graphically, the factored stochastic tree facilitates both the comprehension and the critiquing of a proposed model by others in the medical community. It is also of substantial benefit in formulating a stochastic model, because it allows the analyst to focus formulation effort on one factor of the model at a time, without needing to attend to other irrelevant aspects. In other words, factors are *modular*, in the sense that tolls and rate parameters may be altered in a given factor, or the entire factor may be replaced by one with different structure, with little or no change in the remaining factors of the model. This is the same reason the subtree display is used in Markov-cycle-tree diagrams. The factored stochastic-tree approach therefore confers on stochastic-tree modeling the same benefits as subtrees do for Markov-cycle trees, but in addition inherits the previously known advantages of stochastic-tree models over Markov-cycle trees.

References

1. Hollenberg JP. Markov cycle trees: a new representation for complex Markov processes (abstr). Med Decis Making. 1984;4:529.
2. Beck JR, Kassirer JP, Pauker SG A convenient approximation of life expectancy (the "DEALE"): I. Validation of the method. Am Med. 1982;73:883–8.
3. Hazen GB. Stochastic trees: a new technique for temporal medical decision modeling. Med Decis Making. 1992;12:163–78.
4. Matchar DB, Pauker SG. Transient ischemic attacks in a man with coronary artery disease: two strategies neck and neck. Med Decis Making. 1986;6:239–49.
5. Roach PJ, Fleming C, Hagen MD, Pauker SG. Prostatic cancer in a patient with asymptotic HIV infection: are some lives more equal than others? Med Decis Making. 1988;8:132–44.
6. Hillner BE, Hollenberg JP, Pauker SG. Postmenopausal estrogens in prevention of osteoporosis. Am J Med. 1986;80:1115–27.
7. Tsevat J, Eckman, MH, McNutt RA, Pauker SG. Warfarin for dilated cardiomyopathy: a bloody tough pill to swallow? Med Decis Making. 1989;9:162–9.

Appendix

Error Bounds

It is desirable to obtain error bounds on the deviations $|L_n(y) - L(y)|$ between the calculated quality-adjusted durations $L_n(y)$ and the true values $L(y)$ to which they converge. Bounds on this error may be calculated using the following result.

Theorem 1: For each factor i and product state y, let

$$J_i(y) = [j | \varphi_{ij}(y) \text{ is nonfatal}]$$

that is, $J_i(y)$ is the set of nonfatal transitions from state y in factor i. Let $L(y)$ be the true quality-adjusted duration associated with state y, and suppose $w_i(y)$ and $L_n(y)$ are calculated according to equations 3 and 4 above. Let $B_0(y)$ be an upper bound on $|L(y)|$. If for $n = 1,2,3,\ldots$

$$B_{in}(y) = \frac{\sum_{j \in J_i(y)} \lambda_{ij}(y) B_{n-1}(\varphi_{ij}(y))}{\sum_j \lambda_{ij}(y)} \quad (5)$$

$$B_n(y) = \sum_i w_i(y) B_{in}(y) \quad (6)$$

then

$$|L_n(y) - L(y)| \leq B_n(y)$$

for all n and y. Moreover, if

$$\sum_i \sum_{j \in J_i(y)} \lambda_{ij}(y) > 0$$

(that is, there is some mortality risk from each state y) then $B_n(y)$ converges to zero as $n \to \infty$.

Calculations 5 and 6 can be performed along with the value iteration algorithm given above, and the quantity $\max_y B_n(y)$ may be used as a termination criterion. This is in fact what was done for the calculations of table 1. The initial upper bound $B_0(y)$ on $|L(y)|$ was taken to be $(\mu_o + \mu_c)^{-1} = 3.612$, the mean quality-adjusted lifetime achievable by staying well until death from either cardiomyopathy or background causes. The termination criterion was that maximum error not exceed 0.0005 (accuracy to three decimal places). Termination occurred at $n = 10$ with $\max_y B_n(y) = 1.729 \times 10^{-4}$.