Evaluation quantitative des séquences d'événements par la théorie des langages probabilistes

Dorina-Romina Ionescu, Nicolae Brinzei, Jean-François Pétin

EVALUATION QUANTITATIVE DES SEQUENCES D'EVENEMENTS PAR LA THEORIE DES LANGAGES PROBABILISTES

QUANTITATIVE ASSESSMENT OF EVENTS SEQUENCES BY PROBABILISTIC LANGUAGES THEORY

Ionescu D., Brînzei N., Pétin J-F.
Université de Lorraine, CRAN, UMR 7039
2, avenue de la forêt de Haye, Vandœuvre-lès-Nancy Cedex, 54518, France
E-mail : {dorina-romina.ionescu, nicolae.brinzei, jean-francois.petin}@univ-lorraine.fr

Résumé

Dans les études de sûreté de fonctionnement des systèmes dynamiques et en particulier des systèmes de contrôle-commande, il est nécessaire d'évaluer la probabilité d'occurrence des séquences d'événements qui décrivent l'évolution de ces systèmes ou qui sont considérés critiques. Dans cet article, on utilise les langages probabilistes pour réaliser l'évaluation quantitative de ces séquences. Premièrement, le système à étudier est modélisé par un automate à états finis. L'automate à états finis est ultérieurement transformé dans un automate probabiliste (p-automate) en utilisant la technique des chaînes de Markov à temps discret immergée dans un processus stochastique continu. La détermination formelle des sous-langages associés à chaque état du p-automate permet de calculer leur probabilité d'occurrence et la probabilité d'occurrence des séquences d'événements extraites de ces sous-langages.

Summary

In dependability studies of dynamic systems and particularly in the case of the instrumentation and control systems, it is necessary to assess the probability of events sequences that describe the system evolution or that are considered dangerous. In this paper, the probabilistic languages framework is used in order to realize the quantitative assessment. Firstly, the studied system is modeled by a finite state automaton. This finite state automaton is ulterior transformed in a probabilistic automaton (p-automaton) using the embedded discrete time Markov chain. The formal determination of the sub-languages afferents at each state of the p-automaton enables to calculate their occurrence probability and the occurrence probability for every events sequence that can be subtract from these sub-languages.

Introduction

La plupart des études probabilistes de Sûreté de Fonctionnement, basées en général sur des modèles booléens (arbres de défaillances, diagrammes de fiabilité, etc.), permettent de réaliser :

- une analyse qualitative, telle que la détermination des coupes (sous-ensemble de composants dont leurs défaillances simultanées entraîne la défaillance du système) ;
- une analyse quantitative pour évaluer la probabilité d'occurrence d'un événement redouté ou les indicateurs FMDS, en utilisant soit le théorème de Sylvester-Poincaré en se basant sur les probabilités des coupes, soit la technique de décomposition par des diagrammes de décision binaires (Coudert et Madre, 1992 ; Dutuit et Rauzy, 2005 ; Ibanez-Llano *et al.,* 2010).

L'évaluation de la probabilité d'occurrence d'un événement redouté en utilisant les *coupes* suppose l'Independence totale entre les événements qui composent la coupe. Cependant, cette hypothèse n'est pas adaptée pour les systèmes dynamiques et réparables, dû à la relation d'ordre entre l'occurrence des événements :

- l'occurrence d'un événement peut dépendre de l'occurrence antérieure des autres événements (dans les systèmes dynamiques automatisés, l'occurrence d'un événement peut être interdite par le sous-système de commande en fonction de l'occurrence précédente, ou non, de certains événements).
- l'impact d'une séquence d'événements sur la défaillance du système peut être diffèrent en fonction de l'ordre d'occurrence des événements dans la séquence (par exemple : l'événement e_1 suivi par l'événement e_2 conduit vers un état indésirable tant que e_2 suivi par e_1 peut n'avoir aucun impact).

Dans ces cas, la valeur de la probabilité obtenue en se basant sur le calcul des coupes représente une surévaluation de la probabilité réelle recherchée, car une coupe peut contenir toutes les séquences d'événements appartenant à la coupe dans n'importe quel ordre. De plus, l'approche basée sur les coupes cadre bien aux systèmes non-réparables. Cependant, en fiabilité dynamique il est nécessaire de prendre en considération aussi les changements entre différents modes de fonctionnement/défaillance ; dans ce cas les systèmes doivent être considérées comme réparables et seulement les modèles dynamiques sont capables de prendre en compte les reconfigurations du système. Pour toutes ces raisons l'analyse basée sur les coupes doit être enrichie par la détermination et l'évaluation des séquences d'événements.

Cette évolution justifie l'utilisation des modèles basés sur l'espace d'états pour prendre en compte la dynamique du système et l'impact des défaillance/réparation du composant sur l'état du système. En ce sens, des approches d'analyse quantitative telles que celles proposées par (Papazoglou, 1998 ; Bouissou, 2006) exploitent les séquences d'événements (au lieu des coupes ou des diagrammes de décision binaires) pour déterminer la probabilité d'occurrence d'un événement redouté.

Cet article examine l'utilisation des langages probabilistes définis par (Garg *et al*., 1999) comme un cadre formel pour l'analyse qualitative/quantitative des séquences d'événements.

L'article est structuré dans la manière suivante : la section 2 présente l'état de l'art et met en évidence les limites actuelles des approches existantes pour évaluer les séquences d'événements. La section 3 introduit le cadre formel des langages probabilistes (Garg *et al*., 1999) et les notations utilisées. Le cas d'étude utilisé comme support pour la suite de l'article est

présenté dans la section 4. Dans la section 5 on propose une méthodologie basée sur les langages probabilistes permettant de déterminer et d'évaluer la probabilité d'occurrence des séquences d'événements. La section 6 présente l'application de l'approche proposée au cas d'étude. Enfin la dernière section présente les conclusions de ces travaux et identifie des orientations futures pour nos travaux de recherche.

Etat de l'art

Les approches qui existent dans le domaine de l'évaluation quantitative peuvent être classifié en deux grandes catégories :

- *les modèles booléens* dans lesquels la défaillance du système s'exprime en lien avec les défaillances de ses composants utilisant la fonction de structure booléenne ;
- *les modèles basés sur l'espace d'états* qui décrivent le comportement fonctionnel et dysfonctionnel d'un système par des états et des transitions entre ces états, son évolution est décrite par le langage associé.

1 Les modèles booléens

Parmi les modèles booléens les plus importants on retrouve : les Arbres d'Evénements et les Arbres de Défaillances (classiques/dynamiques). Les *Arbres d'Evénements* (Papazoglou, 1998) sont des modèles qui peuvent être discrétisés, en fonction de la signification ou des effets possibles des événements pouvant se produire, dans une série des événements simples. Les *Arbres d'Evénements* sont adaptés à la modélisation et l'évaluation des séquences d'événements dans les cas des systèmes non-réparables tel que les systèmes de protection et de sûreté. Par contre, les systèmes réparables ou les systèmes de contrôle-commande ne peuvent pas être modélisés par les *Arbres d'Evénements*. Les *Arbres de Défaillances* sont construits en fonction d'un événement indésirable donné qui est décomposé en événements de base jusqu'à ce que cette décomposition soit impossible ou jugée inutile Ces modèles sont efficaces pour l'analyse basée sur les coupes ou les diagrammes de décision binaires (Coudert et Madre, 1992 ; Rauzy, 1993 ; Dutuit et Rauzy, 2005) pour des systèmes dont on ne considère pas leur évolution dans le temps.

2 Les modèles basés sur l'espace d'états

Les approches basées sur l'espace d'états (Cassandras et Lafortune, 2008) les plus couramment utilisées dans les études de sûreté de fonctionnement sont : les automates à états finis et la théorie des langages, les réseaux de Petri et les chaînes de Markov. Dans le contexte de cet article relativement à l'évaluation probabiliste en se basant sur des séquences d'événements, on peut mentionner aussi d'autres modèles. Les BDMP (*Boolean logic Driven Markov Process)* développé par Bouissou et Bon (2003) remplace les événements de base d'un arbre de défaillance par des chaînes de Markov ; cette combinaison (arbre des défaillances et chaînes de Markov) présente des caractéristiques dynamiques pour couvrir l'ordre et l'impact des occurrences des événements. Les séquences d'événements sont déterminées par l'exploration du modèle à partir de l'état initial du système (Bouissou, 2006). Les *Automates Stochastiques Hybrides* sont des automates stochastiques à états finis temporisés avec des équations différentielles associées à chaque état (Perez-Castaneda et al., 2011). Ce modèle permet de déterminer les séquences d'événements et leur fréquences d'occurrence par la simulation de Monte-Carlo (Aubry et al., 2012 ; Babykina et al. 2012), mais il n'est pas possible d'obtenir une solution analytique.
Tous ces modèles peuvent être utilisés, selon leur caractéristiques stochastiques ou déterministes pour l'analyse des séquences d'événements (Bouissou, 2006 ; Chaux *et al*., 2013), pour résoudre deux types de problèmes : la *détermination des séquences* et l'*évaluation de leur probabilité d'occurrence*. Pour déterminer les séquences d'événements, Chaux et al. (2013) proposent d'utiliser la théorie des langages déterministes pour calculer l'ensemble minimal de séquences qui vise à reconstruire le langage défaillant. Les calculs sont possibles seulement pour les systèmes cohérents et, dû au fait qu'il retient seulement un ensemble minimal de séquences, cette approche ne permet plus de faire une quantification probabiliste pour les séquences déterminées. Le concept de séquence minimale est également pris en compte dans (Bouissou, 2006) et, sa définition ainsi que l'obtention des séquences d'événements par exploration d'un modèle BDMP ou réseau de Petri permettent également de réaliser une quantification probabiliste des événements dangereux ou des états dans lesquels ces séquences amènent le système.

Concernant l'évaluation de la probabilité des séquences, les approches existantes présentent quelques limitations :

- les modèles basés sur l'espace d'états comme les chaînes de Markov sont efficaces pour déterminer les probabilités d'état du système. Par conséquent, une séquence d'événements sera assimilée à son état final qui doit être considéré comme état absorbant. Dans ce cas, la probabilité de cet état représente la probabilité du sous-langage (toutes les séquences ayant comme cible cet état) qui permet au système d'atteindre cet état.
- le calcul de la probabilité d'une séquence est basé sur la probabilité d'occurrence des événements appartenant à la séquence. Dans les approches existantes, la valeur de la probabilité d'un événement reste la même quel que soit l'état où l'événement se produit. Cependant, dans les systèmes dynamiques un événement donné peut avoir différentes probabilités d'occurrence en fonction de l'état du système dans lequel cet événement a lieu, même si son signification – défaillance ou réparation du composant c_i – reste la même dans toutes les séquences.

Les *langages probabilistes* (Garg *et al*., 1999), qui appartient également à la classe des approches basées sur l'espace d'états et qui combinent les avantages de la théorie des langages et de la théorie des probabilités, constitue un cadre formel intéressant pour l'analyse des séquences d'événements critiques. La section suivante présente la théorie des *langages probabilistes*.

Les langages probabilistes

La théorie des langages probabilistes a été développée par Garg et al., (1999) afin de modéliser le comportement des systèmes à événements discrèts (SED). Afin de simplifier du point de vue mathématique la définition formelle d'un langage probabiliste, un événement particulier appelé « événement de terminaison » noté e_Δ est introduit pour représenter le fait que l'état du système obtenu après l'occurrence d'une séquence est un état terminal pour l'étude. Un état du système est terminal s'il représente un intérêt particulier pour l'étude (e.g. la réalisation de la mission du système, un état de défaillance dangereux). Ainsi, le comportement du système est donné par l'ensemble Ω de toutes les séquences de longueur finie suivies ou non par l'événement de terminaison :

$$\Omega = \Sigma^*(e_\Delta + \varepsilon) = \Sigma^* e_\Delta \cup \Sigma^* \quad \{1\}$$

Dans cette équation Σ représente l'ensemble de tous les événements, appelé alphabet, Σ^* est l'ensemble des toutes les séquences d'événements de longueur finie de l'alphabet Σ (* est l'opération d'itération appelé aussi « fermeture de Kleene ») et ε représente une séquence d'événements vide.

Pour déterminer la probabilité d'occurrence d'une séquence d'événements *s*, il faut considérer non pas seulement l'occurrence indépendante de cette séquence *s*, mais aussi l'occurrence de toutes les séquences qui démarrent par *s* (toutes les séquences ayant le préfixe s). L'ensemble de toutes les séquences ayant une séquence donnée *s* comme préfixe est donné par :

$$< s >= \{st | \, st \in \Omega\} \quad \{2\}$$

Définition 1 (Garg *et al.,* 1999) : Soit l'espace mesurable (Ω, F), ou $\Omega = \Sigma^*(\mathrm{e}_\Delta + \varepsilon)$ et *F* est une σ-algèbre générée par $\{< s > | \, \mathrm{s} \in \Omega\}$. Un *langage probabiliste* (ou *p-langage*) L est une mesure de probabilité sur l'espace mesurable (Ω, F). Ainsi, la *probabilité de terminaison d'une séquence s* (probabilité que l'évolution d'un système se termine après l'occurrence de *s*) est donnée par la relation suivante :

$$\mathbb{P}(se_\Delta) = \mathbb{P}(s) - \sum_{e\in\Sigma} \mathbb{P}(se), \forall s \in \Sigma^* \quad \{3\}$$

où $\mathbb{P}(s)$ représente la probabilité d'occurrence de *s* et $\mathbb{P}(se)$ représente la probabilité que le système continue à évoluer au-delà de s.

On peut noter que les probabilités de terminaison correspondant à des séquences distinctes sont mutuellement exclusives. Ainsi, la probabilité cumulée que le système arrive dans un état terminal peut être obtenue en ajoutant les probabilités individuelles de toutes les séquences possibles :

$$\mathbb{P}(\mathrm{Syst}\, e_\Delta) = \sum_{s\in\Sigma^*} \mathbb{P}(se_\Delta) \quad \{4\}$$

Des fois c'est plus facile à décrire un *p-langage* en utilisant un automate associé qui peut reconnaître ce langage.

Définition 2 : Un automate probabiliste (ou *p-automate*) sur un alphabet (ensemble d'événements) Σ est défini par le quintuple suivant :

$$A_p = (\mathrm{X}, \Sigma, f, \mathbb{P}, x_0) \quad \{5\}$$

où :

- X est un ensemble fini d'états ;
- Σ est un ensemble fini d'événements appelé alphabet ;
- $f: \mathrm{X} \times \Sigma \to X$ est la fonction de transition qui associe à chaque état de départ et à chaque événement un état d'arrivée ;
- $\mathbb{P}: \mathrm{X} \times \Sigma \times \mathrm{X} \to [0, 1]$ est la fonction de probabilité de transition affectant à chaque transition une probabilité d'occurrence qui vérifie la relation suivante :

$$\sum_{x_j\in X} \sum_{e\in\Sigma} \mathbb{P}(x_i, e, x_j) \leq 1, \forall x_i \in X \quad \{6\}$$

- x_0 représente l'état initial.

L'évolution d'un p-automate est la suivante : si le système est dans l'état x_i, l'occurrence de l'événement e permet de franchir la transition vers l'état x_j avec la probabilité $\mathbb{P}(x_i, e, x_j)$. La fonction de probabilité de transition peut être étendue à des chemins $\pi \subset \mathrm{X}(\Sigma\mathrm{X})^*$ dans le p-automate A_p (un chemin étant obtenu par la concaténation des transitions, où l'état d'arrivée et respectivement l'état de départ de deux transitions consécutives coïncident). La probabilité des chemins est définie par l'équation suivante :

$$x_j \in X : \mathbb{P}(\pi e x_j) = \mathbb{P}(\pi)\mathbb{P}(x_{|\pi|} e x_j) \quad \{7\}$$

Chaque p-automate définit un p-langage et inversement chaque p-langage peut être représenté par un p-automate. On peut remarquer que la probabilité d'une séquence $\mathbb{P}(s)$ dans l'équation (3) est donnée par la somme des probabilités de tous les chemins qui démarrent par s données par l'équation (7).

Définition 3 (Garg *et al.,* 1999) : Etant donné un p-langage L, le p-automate $A_p = (\mathrm{X}, \Sigma, f, \mathbb{P}, x_0)$ avec :

$$\forall s, t \in \Sigma^*, e \in \Sigma : \mathbb{P}(s, e, t) = \begin{cases} \frac{L(t)}{L(s)}, & \text{si } \mathrm{t} = se \\ 0, & \text{autrement}. \end{cases} \quad \{8\}$$

génère le p-langage L.

Les données d'entrée pour cette approche basée sur l'espace d'état sont exprimées sous la forme d'un ou plusieurs langages probabilistes $L(s), \forall s \in \Sigma^*$. Ceci signifie, qu'apriori, on est capable de donner pour un système l'ensemble de séquences qu'il peut suivre et que, pour chaque séquence, on sera capable de donner la probabilité d'occurrence de toutes les séquences.

Présentation du cas d'étude

Afin d'illustrer l'application de la théorie des langages probabilistes, nous considérons un système de commande de la température d'un four (Figure 1). Il fonctionne de la manière suivante : au démarrage la température du four est commandée par un régulateur proportionnel-intégral (PI) qui a le rôle de maintenir la température à une valeur de consigné donnée. Lorsque ce régulateur tombe en panne (avec un taux de défaillance λ_{PI}), la commande de la température sera assuré par un régulateur de type « tout ou rien » (TOR) qui devra maintenir la température dans une plage donnée. Le régulateur TOR peut à son tour tomber en panne avec un taux de défaillance λ_{TOR}. Lorsque l'un ou l'autre des régulateurs est défaillant il sera réparé, le processus de réparation étant caractérisé par un taux de réparation μ_{PI}, respectivement μ_{TOR}. Ce fonctionnement représente le fonctionnement simplifié du cas d'étude défini en (Perez *et al*., 2011) afin d'illustrer les problèmes posés en fiabilité dynamique. Le modèle présenté ici a été obtenu en éliminant la partie continue du système décrivant l'évolution de la température (décrite par des équations différentielles) dans les différents états discrets, ainsi que le système de diagnostic des défaillances basé sur le franchissement des seuils de la variable « température ».

Les valeurs des différentes taux sont les suivantes : $\lambda_{PI} = 3.5 \cdot 10^{-5}\, h^{-1}$, $\lambda_{TOR} = 2 \cdot 10^{-5} h^{-1}$, $\mu_{PI} = 8 \cdot 10^{-2} h^{-1}$, $\mu_{TOR} = 10 \cdot 10^{-2} h^{-1}$ et l'automate à états finis qui modélise le système est présenté dans la Figure 2. Cet automate représente une chaîne de Markov à temps continu.

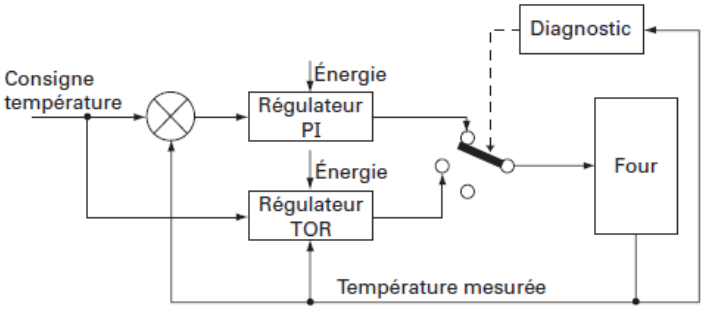

Figure 1. Schéma-bloc du système de commande de la température d'un four

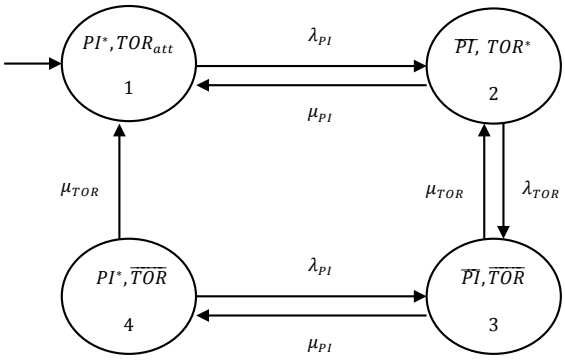

Figure 2. La chaîne de Markov à temps continu du cas d'étude,
où X^* représente le régulateur X $(PI\ ou\ TOR)$ qui est utilisé pour la commande du four,
X_{att} représente le régulateur X en état d'attente et $\overline{X}$ représente le régulateur X en panne.

Approche proposée

L'approche que nous proposons pour déterminer de manière formelle l'ensemble de toutes les séquences d'événements et évaluer leur probabilité en utilisant les langages probabilistes est constituée des quatre étapes et elle est présentée schématiquement dans la Figure 3.

1 Etape 0 : modélisation du système

Si on souhaite utiliser la théorie des langages probabilistes telle quelle a été proposée par Garg *et al.* (1999), on devait être capable de fournir par l'ensemble de séquences de longueur finie Σ^* (cet ensemble pourra être infini) et, pour chaque séquence $s \in \Sigma^*$, on devait donner sa probabilité d'occurrence $\mathbb{P}(s)$ incluant l'occurrence de s, mais aussi de toutes les séquences ayant s comme préfixe. Ces données représentent le langage probabiliste conformément à la définition 1. Cette définition et cette manière de poser le problème sont appropriées aux études concernant la conception de la commande par supervision (Kumar et Garg, 2001). Des nombreux autres travaux existent dans la littérature concernant l'utilisation des langages probabilistes pour la conception de la commande par supervision.

Dans les études de sûreté de fonctionnement, il sera difficile, voire impossible, de procéder de cette manière et d'exprimer d'abord l'ensemble de séquences et leurs probabilités. En sûreté de fonctionnement, on se trouve plutôt dans la situation inverse : on cherche à déterminer les différentes séquences $s \in \Sigma^*$ qui présentent un intérêt pour l'étude et, ensuite, d'évaluer leur probabilité d'occurrence, au sens de leur probabilité de terminaison, c'est-à-dire la probabilité d'occurrence exclusive d'une séquence $s \in \Sigma^*$ sans la prise en compte des autres séquences ayant s comme préfixe.

Afin de pallier ces inconvénients, notre approche consiste d'abord à modéliser le système comme un automate à états finis et non comme un langage probabiliste donnant apriori les probabilités d'occurrence d'un ensemble de séquences. Pour le cas d'étude considéré, cet automate est celui donné par la figure 2.

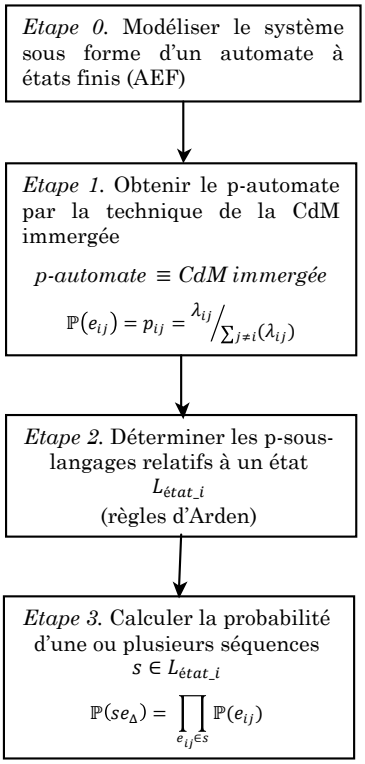

Figure 3. Approche proposée pour l'évaluation probabiliste des séquences d'événements

2 Etape 1 : détermination des probabilités d'événements

L'automate à états finis représentant le système doit être accompagné des probabilités discrètes d'occurrence de chaque transition pour pouvoir le qualifier comme un p-automate (conformément à la définition 2). Puisque dans les études de sûreté de fonctionnement les lois des probabilités décrivant les phénomènes aléatoires sont plutôt des lois continues dans le temps, pour déterminer les probabilités discrètes requises par la théorie de p-automates nous proposons d'utiliser *la chaîne de Markov à temps discret immergée* (CdM immergée) dans un processus stochastique continu.

Cette chaîne de Markov immergée est obtenue en considérant, dans le processus stochastique continu, les instants de saut à la fin du temps de séjour dans l'état courant et elle permet au système d'atteindre la distribution stationnaire des probabilités d'état. Par conséquent la probabilité p_{ij} d'un événement e_{ij} associé à une transition qui part d'un état donné s_i vers un autre état s_j est déterminée par le ratio entre le taux de cette transition et la somme des taux pour toutes les transitions sortantes de l'état s_i. Cette probabilité est donnée par la relation suivante :

$$\mathbb{P}(e_{ij}) = p_{ij} = {\lambda_{ij}}/{\sum_{j\neq i}(\lambda_{ij})} \quad \{9\}$$

La chaîne de Markov immergée du cas d'étude est présentée dans la figure 4.

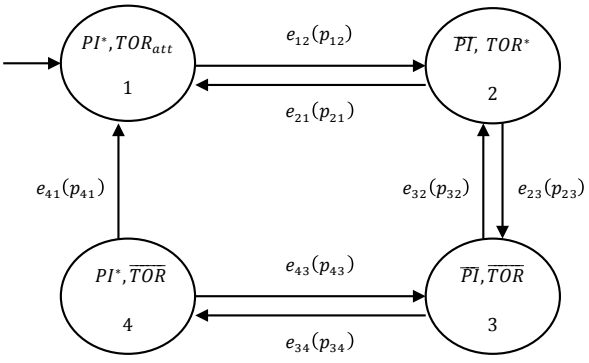

Figure 4. La chaîne de Markov à temps discret qui représente le p-automate du cas d'étude,
où e_{ij} représente l'événement dont son occurrence amène le système depuis l'état s_i dans l'état s_j et p_{ij} représente la probabilité de transition depuis l'état s_i dans l'état s_j

Par exemple pour le cas d'étude les probabilités d'occurrences des événements faisant sortir le système de l'état 2 sont les suivantes :

$$\mathbb{P}(e_{21}) = p_{21} = \frac{\mu_{PI}}{\lambda_{TOR} + \mu_{PI}}$$

$$\mathbb{P}(e_{23}) = p_{23} = \frac{\lambda_{TOR}}{\lambda_{TOR} + \mu_{PI}}$$

Les probabilités d'occurrence des événements faisant sortir le système de l'état 3 sont les suivantes :

$$\mathbb{P}(e_{32}) = p_{32} = \frac{\mu_{TOR}}{\mu_{TOR} + \mu_{PI}}$$

$$\mathbb{P}(e_{34}) = p_{34} = \frac{\mu_{PI}}{\mu_{TOR} + \mu_{PI}}$$

On note que l'événement représentant la réparation du régulateur PI a des probabilités d'occurrence différentes p_{21} et, respectivement, p_{34} dépendant de l'état dans lequel le système se trouve lorsque l'événement a lieu.

On remarque également que la somme de toutes les transitions qui partent d'un état donné est égale à 1.

3 Etape 2 : détermination des sous-langages

Cette étape permet de déterminer formellement l'ensemble de séquences que le système peut suivre par la détermination des sous-langages associés à chaque état du système.

Pour montrer l'application de la théorie des langages probabilistes pour le calcul de la probabilité des séquences, nous allons nous intéresser à la séquence d'événements qui atteignent l'état 3. Cet état représente l'état dangereux pour le four car, les deux régulateurs étant défaillants, la température n'est plus contrôlée. Afin d'utiliser l'équation (3), qui donne l'expression de cette probabilité, l'état 3 doit être considéré comme état terminal. Ainsi pour une séquence s qui termine dans l'état 3, en considérant premièrement que la probabilité d'occurrence d'une séquence est égale au produit des probabilités des transitions (équation 7) et en considérant les événements e_{32} et e_{34} qui permettent au système de continuer son évolution au-delà de l'état 3, la probabilité de terminaison de la séquence s sera :

$$\mathbb{P}(se_{\Delta}) = \mathbb{P}(s) - \sum_{e \in \Sigma} \mathbb{P}(se) = \mathbb{P}(s) - \left(\mathbb{P}(se_{32}) + \mathbb{P}(se_{34})\right)$$

$$= \mathbb{P}(s) - \left(\mathbb{P}(s)\mathbb{P}(e_{32}) + \mathbb{P}(s)\mathbb{P}(e_{34})\right) = \mathbb{P}(s) - \mathbb{P}(s)\left(\mathbb{P}(e_{32}) + \mathbb{P}(e_{34})\right)$$

$$= \mathbb{P}(s) - \mathbb{P}(s)(p_{32} + p_{34}) = \mathbb{P}(s) - \mathbb{P}(s) \cdot 1 = 0$$

Cette approche a été proposée dans la théorie des p-langages et a été développée dans le cadre de la synthèse de la stratégie de commande par supervision (Wang et Ray, 2004). Le résultat obtenu montre que, dans les études de sûreté de fonctionnement, les calculs des probabilités des séquences d'événements ne sont pas possibles si le p-langage est couplé avec la chaîne de Markov immergée sans certaines précautions (si l'état terminal d'une séquence n'est pas un état absorbant, sa probabilité est toujours égale à 0).

Pour éviter ceci, nous proposons de déterminer l'ensemble de toutes les séquences d'événements du system, représenté par le langage L_{Sys}. Ce langage peut être représenté comme la réunion des sous-langages associés aux états du système. Un sous-langage L_i associé à un état s_i est défini comme l'ensemble de toutes les séquences d'événements, qui conduisent le système à partir de son état initial dans l'état considéré s_i.

$$L_{Sys} = \bigcup_{s_i} L_i \qquad \{10\}$$

Pour déterminer les sous-langages L_i on propose d'utiliser la théorie des langages rationnels et plus précisément le lemme d'Arden (Carton, 2008).

Lemme d'Arden : Soit deux langages *A* et *B* et soit l'équation :

$$L_i = L_i A + B \qquad \{11\}$$

où L_i représente le langage inconnu :

1. si $\varepsilon \notin A$ la solution unique de l'équation (11) est $L_i = BA^*$.
2. si $\varepsilon \in A$ les solutions ont la forme $L_i = (B + C)A^*$ où $C \subseteq \Sigma^*$.

Ce lemme est utilisé surtout dans le cas ou $\varepsilon \notin A$ et le langage $L_i = BA^*$ représente dans ce cas la solution unique de l'équation (11).

Pour les états s_i du système et en considérant les séquences à partir de tous les autres états $s_j \neq s_i$ et arrivant dans l'état considéré s_i en une seule transition, on obtient un ensemble de *n* équations (11) (où *n* est le nombre d'états du système). Ces équations permettent d'obtenir tous les sous-langages L_i.

Ainsi, toutes les séquences d'événements d'un système sont déterminées formellement sans faire appel à l'exploration du modèle.

4 Etape 3 : calcul des probabilités des séquences

Chaque séquence d'événements qu'un système peut suivre, $s = e_{12}e_{23} \dots e_{(n-1)n}$, et qui commence d'un état initial s_1 pour arriver dans un état s_n peut être extraite du langage L_{Sys} obtenu à l'étape précédente. En utilisant la probabilité de chaque transition du p-automate obtenu à l'étape 1 nous proposons d'utiliser l'équation suivante afin d'obtenir la probabilité de terminaison pour la séquence $s = e_{12}e_{23} \dots e_{(n-1)n}$:

$$\mathbb{P}(se_{\Delta}) = \prod_{e_{ij} \in s} \mathbb{P}(e_{ij}),\ \forall s \in \Sigma^* \quad \{12\}$$

Cette équation permet de calculer la probabilité d'occurrence d'une séquence comme le produit des probabilités de tous ses événements e_{ij}.

Application au cas d'étude

1 Application de l'approche proposée

1.1 Etape 0 : L'automate à états finis qui modélise le système de commande de la température du four est une chaîne de Markov à temps continu et il est présenté dans la figure 2.

1.2 Etape 1 : Le p-automate représenté par la chaîne de Markov à temps discret immergée dans la chaîne de Markov à temps continu est présenté dans la figure 4. Les probabilités des événements associés aux transitions p_{ij} sont calculées en utilisant l'équation (9).

1.3 Etape 2 : Les sous-langages associés à chaque état représentent les sous-ensembles de séquences qui amènent le système depuis l'état initial dans chacun de ses états. Si on considère chaque état de l'automate comme état terminal (non absorbant) on obtient les équations suivantes :

$$L_1 = L_2 e_{21} + L_4 e_{41} \quad \{13\}$$

$$L_2 = L_1 e_{12} + L_3 e_{32} \quad \{14\}$$

$$L_3 = L_2 e_{23} + L_4 e_{43} \quad \{15\}$$

$$L_4 = L_3 e_{34} \quad \{16\}$$

Par exemple, la première équation (13) donne le sous-langage affèrent à l'état s_1 qui contient toutes les séquences d'événements amenant le système dans cet état en une seule transition :

- à partir de l'état s_2 (toutes les séquences d'événements qui ont s_2 comme état terminal : L_2) en franchissant la transition de s_2 à s_1 sous l'occurrence de l'événement e_{21} représentant la réparation du régulateur PI ;
- ou à partir de l'état s_4 (toutes les séquences d'événements qui ont s_4 comme état terminal : L_4) en franchissant la transition de s_4 à s_1 sous l'occurrence de l'événement e_{41} représentant la réparation du régulateur TOR.

En utilisant le lemme d'Arden les équations (13-16) donnent les expressions suivantes pour les sous-langages :

- pour l'état 1 :

$$L_1 = [e_{12}(e_{23}(e_{34}e_{43})^* e_{32})^*(e_{21} + e_{23}(e_{34}e_{43})^* e_{34}e_{41})]^* \quad \{17\}$$

- pour l'état 2 :

$$L_2 = L_1 e_{12}[e_{23}(e_{34}e_{43})^* e_{32}]^* = [e_{12}(e_{23}(e_{34}e_{43})^* e_{32})^*(e_{21} + e_{23}(e_{34}e_{43})^* e_{34}e_{41})]^* e_{12}[e_{23}(e_{34}e_{43})^* e_{32}]^* \quad \{18\}$$

- pour l'état 3 :

$$L_3 = L_2 e_{23}(e_{34}e_{43})^* = [e_{12}(e_{23}(e_{34}e_{43})^* e_{32})^*(e_{21} + e_{23}(e_{34}e_{43})^* e_{34}e_{41})]^* e_{12}[e_{23}(e_{34}e_{43})^* e_{32}]^* e_{23}(e_{34}e_{43})^* \quad \{19\}$$

- pour l'état 4 :

$$L_4 = L_3 e_{34} = [e_{12}(e_{23}(e_{34}e_{43})^* e_{32})^*(e_{21} + e_{23}(e_{34}e_{43})^* e_{34}e_{41})]^* e_{12}[e_{23}(e_{34}e_{43})^* e_{32}]^* e_{23}(e_{34}e_{43})^* e_{34} \quad \{20\}$$

1.4 Etape 3 : A partir des expressions des sous-langages, il est possible :

- de déterminer la probabilité de chaque sous-langage, c'est-à-dire la probabilité de toutes les séquences amenant le système depuis l'état initial dans chacun de ses états ;
- d'extraire des séquences ayant un intérêt particulier pour l'analyse de sûreté de fonctionnement du système étudié et d'en calculer sa probabilité.

Nous allons illustrer ces calculs pour deux sous-langages : L_1 et L_3. Le langage L_1 représente toutes les séquences possibles qui amènent le système dans son état de bon fonctionnement. Le langage L_3 représente toutes les séquences possibles qui amènent le système dans son état dangereux où la température du four n'est plus contrôlée.

Pour simplifier on remarque que $p_{12} = 1$ (car il y a une seule transition sortante de l'état s_1) et on utilisera les notations suivantes : $p_{23} = p$; $p_{21} = (1-p)$; $p_{34} = p'$; $p_{32} = (1-p')$; $p_{41} = p''$; $p_{43} = (1-p'')$.

1.4.1 Evaluation des probabilités pour le sous-langage L_1

En utilisant l'équation (17) on calcule les probabilités des différentes itérations (opération *) qui composent l'expression régulière du sous-langage L_1 et on obtient la relation suivante :

$$\mathbb{P}(L_1) = \sum_{s \in \Sigma^*} \mathbb{P}(se_{\Delta}) = \frac{(1-p)[1-p'(1-p'')]+pp'p''}{1-p'(1-p'')-p(1-p')} = 1 \quad \{21\}$$

Le résultat obtenu pour la probabilité de sous-langage L_1 est celui qui est attendu, parce que pour le p-automate de la Figure 4 *la probabilité de toutes les séquences qui reviennent à l'état s_1 en partant de de lui-même* doit être égale à 1, lorsque le système ne s'arrête pas dans un autre état intermédiaire (s_2, s_3 ou s_4) ou qu'il n'y a pas des séquences partant de s_1 et pouvant amener le

système dans un état (ou sous-ensemble d'états) irréversible à partir d'où le système ne peut plus revenir dans l'état s_1. Ce résultat signifie que le système finira toujours par revenir dans l'état de bon fonctionnement.

Le sous-langage L_1 permet d'extraire de son expression générale, donnée par l'équation (17), des séquences individuelles dont leur probabilité d'occurrence est obtenue en utilisant l'équation (12). Ces probabilités sont présentées dans le tableau 1 pour quelques séquences d'événements.

Séquence (s_i)	**Probabilité de la séquence ($\mathbb{P}(s_i)$)**
$s_1 = e_{01}e_{12}e_{21}$	0.8936
$s_2 = e_{01}e_{12}e_{23}e_{33}e_{32}e_{21}$	$4.8462 \cdot 10^{-6}$
$s_3 = e_{01}e_{12}e_{22}e_{23}e_{32}e_{23}e_{32}e_{21}$	$4.6638 \cdot 10^{-10}$
$s_4 = e_{01}e_{12}e_{23}e_{34}e_{43}e_{33}e_{32}e_{21}$	$7.3098 \cdot 10^{-10}$
$s_5 = e_{01}e_{12}e_{23}e_{34}e_{43}e_{34}e_{43}e_{32}e_{21}$	$2.6820 \cdot 10^{-12}$
$s_6 = e_{01}e_{12}e_{23}e_{34}e_{41}$	$9.9258 \cdot 10^{-5}$
$s_7 = e_{01}e_{12}e_{22}e_{23}e_{33}e_{34}e_{41}$	$1.2239 \cdot 10^{-7}$
$s_8 = e_{01}e_{12}e_{23}e_{33}e_{32}e_{23}e_{34}e_{41}$	$5.3828 \cdot 10^{-10}$
$s_9 = e_{01}e_{12}e_{22}e_{23}e_{32}e_{23}e_{32}e_{23}e_{34}e_{43}e_{34}e_{41}$	$7.8135 \cdot 10^{-18}$
$s_{10} = e_{01}e_{12}e_{21}e_{12}e_{22}e_{23}e_{34}e_{41}$	$2.7426 \cdot 10^{-6}$
$s_{11} = e_{01}e_{12}e_{21}e_{12}e_{23}e_{33}e_{32}e_{21}$	$4.4646 \cdot 10^{-6}$
$s_{12} = e_{01}e_{12}e_{22}e_{21}e_{12}e_{23}e_{34}e_{43}e_{34}e_{41}$	$4.1368 \cdot 10^{-10}$
	$\mathbb{P}(L_1) \cong \sum_{s_i} \mathbb{P}(s_i) \cong 0.8937$

Table 1. Probabilités d'occurrence des quelques séquences qui amènent le système dans l'état 1

Premièrement on constate que la somme des probabilités des séquences d'événements considérées est approximativement égale à 1 et ceci correspond à la solution analytique de la probabilité du sous-langage $\mathbb{P}(L_1)$ donnée par l'équation (21). Egalement on peut identifier les séquences d'événements les plus significatives ayant les plus grandes valeurs pour leur probabilité d'occurrence. De plus, il sera possible également d'estimer l'erreur de calcul due à la non prise en compte des certaines séquences.

1.4.2 Evaluation des probabilités pour le sous-langage L_3

Le sous-langage L_3 décrit une infinité des séquences qui amènent le système dans l'état défaillant s_3 à partir de l'état initial s_1. Sa probabilité d'occurrence est donnée par la relation suivante :

$$\mathbb{P}(L_3) = \frac{p}{1-p'(1-p'')-p(1-p')} \quad \{22\}$$

En faisant les calculs pour quelques séquences nous obtenons leurs probabilités présentées dans le tableau 2.

Séquence (s_i)	**Probabilité de la séquence ($\mathbb{P}(s_i)$)**
$s_1 = e_{12}e_{23}$	$2.4990 \cdot 10^{-4}$
$s_2 = e_{12}e_{23}e_{32}e_{23}$	$3.4700 \cdot 10^{-8}$
$s_3 = e_{12}e_{21}e_{12}e_{23}$	$2.4990 \cdot 10^{-4}$
$s_4 = e_{12}e_{23}e_{34}e_{43}$	$3.8870 \cdot 10^{-8}$
$s_5 = e_{12}e_{23}e_{32}e_{23}e_{34}e_{43}$	$5.3970 \cdot 10^{-12}$
$s_6 = e_{12}e_{21}e_{12}e_{23}e_{34}e_{43}$	$3.8860 \cdot 10^{-8}$
	$\mathbb{P}(L_3) \cong \sum_{s_i} \mathbb{P}(s_i) = 4.9990 \cdot 10^{-4}$

Table 2. Probabilités d'occurrence des quelques séquences qui amènent le système dans l'état 3

L'état 3 représente un état dangereux à cause du fait que la température n'est plus contrôlée et ainsi nous pouvons identifier les séquences d'événements les plus critiques (avec les plus grande valeurs de leur probabilité) du système.

2 Validation analytique des résultats

Afin de prouver la validité des résultats (probabilités des séquences) obtenus, nous proposons l'approche suivante en deux étapes :

1. Déterminer la distribution stationnaire des probabilités d'états pour la chaîne de Markov immergée : $[\pi_1\ \pi_2\ \pi_3\ \pi_4]$ en utilisant les équations :

$$\begin{aligned} \pi &= \pi \cdot M \\ \pi \cdot \overline{1} &= 1 \end{aligned} \quad \{23\}$$

où le vecteur π représente la distribution des probabilités d'état, *M* est la matrice de probabilités de transition et $\overline{1}$ est un vecteur de sommation avec tous ses éléments égal à 1.

En utilisant les équations (23) on détermine l'expression analytique des probabilités d'états :

$\pi_1 = \frac{1-p-p'+pp'+p'p''}{2(1-p'+pp'+p'p'')}$; $\pi_2 = \frac{1-p'+p'p''}{2(1-p'+pp'+p'p'')}$; $\pi_3 = \frac{p}{2(1-p'+pp'+p'p'')}$; $\pi_4 = \frac{pp'}{2(1-p'+pp'+p'p'')}$

2. Prouver que les probabilités de différents sous-langages, obtenues dans le paragraphe antérieur, vérifient la relation suivante :

$$\mathbb{P}(s_i) \cdot \sum_{\substack{s \in \Sigma^* \\ i \xrightarrow{s} j}} \mathbb{P}(se_\Delta) = \mathbb{P}(s_i) \cdot \mathbb{P}\left(L_{i \to j}\right) = \mathbb{P}\left(s_j\right) \quad \{24\}$$

La signification de cette équation est la suivante : la probabilité d'arrivée dans l'état s_j est égale au produit de la probabilité de l'état initial s_i et de la probabilité du sous-langage entre les états s_i et s_j.

- Pour le langage L_1 : $\pi_1 \cdot \mathbb{P}(L_1) = \frac{1-p-p'+pp'+p'p''}{2(1-p'+pp'+p'p'')} \cdot 1 = \pi_1$
- Pour le langage L_3 : $\pi_1 \cdot \mathbb{P}(L_3) = \frac{p}{2(1-p'+pp'+p'p'')} = \pi_3$

Conclusions et perspectives

Dans cet article, nous avons présenté une approche de quantification des séquences d'événements en sûreté de fonctionnement basée sur la théorie des langages probabilistes. Les systèmes étudiés sont modélisés sous forme d'automates à états finis, plus précisément sous forme d'automates probabilistes. Dans ce but, nous proposons d'utiliser une chaîne de Markov à temps discret immergée dans un processus stochastique à temps continu. L'obtention d'une chaîne de Markov à temps discret immergée pour déterminer les probabilités des évènements est possible lorsque le système admet un régime permanent caractérisé par une distribution stationnaire des probabilités d'état. La définition d'une chaîne de Markov à temps discret immergée n'est pas restrictive aux chaînes de Markov à temps continu. Une chaîne de Markov à temps discret immergée peut être définie également pour des processus semi-markoviens ou, sous quelques hypothèses, pour des processus stochastiques plus généraux. Un autre avantage lié à l'utilisation des chaînes de Markov immergées est que celles-ci permettent de déterminer la probabilité d'occurrence d'un événement en fonction de l'état dans lequel le système se trouve (ainsi, le même événement, défaillance ou réparation, peut être caractérisé par des probabilités d'occurrence différentes pendant l'évolution du système). Le lemme d'Arden permet de déterminer l'ensemble de toutes les séquences d'événements, sans explorer le modèle du système. Finalement les probabilités d'occurrence d'un sous-langage et des séquences d'événements pertinentes (pour le bon fonctionnement du système) ou critiques (relativement aux états dangereux) sont déterminées. En conclusion, par cette approche, il est possible d'obtenir formellement l'ensemble des séquences qu'un système peut suivre pendant son évolution et, de calculer analytiquement la probabilité d'occurrence de ces séquences, permettant ainsi de réduire les erreurs des évaluations classiques de ces probabilités ou de quantifier les erreurs dues à la non prise en compte des certaines séquences.

Les travaux que nous avons présentés dans cet article conduisent à plusieurs axes de recherche. Premièrement, le régime de fonctionnement transitoire du système avant l'atteinte d'un régime permanent nécessite de développer une approche de calcul des probabilités d'événements correspondant à ce régime. Nous sommes intéressés, également, à appliquer l'approche proposée à l'étude des systèmes plus complexes intégrant le vieillissement des composants. Enfin, nous nous proposons d'étendre cette approche au contexte de la fiabilité dynamique par la prise en compte de manière explicite de l'évolution des variables physiques (température, pression etc.) décrites par des équations algébro-différentielles associées aux états discrets du modèle.

Références

Aubry J. F., Babykina G., Brinzei N., Medjaher S., Barros A., Bérenguer C., Grall A., Langeron Y., Ngoc Nguyen D., Deleuze G., De Saporta B., Dufour F., Zhang H., "The APPRODYN project: dynamic reliability approaches to modeling critical systems", in Supervision and Safety of Complex Systems , Editors: Matta N., Vandenboomgaerde Y., Arlat J., ISBN 978-1-84821-413-2, Wiley- ISTE, London), 2012, 141-179.

Babykina G., Brinzei N., Aubry J. F., Deleuze G., 2012, Modélisation des systèmes complexes critiques en fiabilité dynamique par automates stochastiques hybrides, évaluation de leur comportement, *18ème Congrès de Maîtrise des Risques et de Sûreté de Fonctionnement, Lambda-Mu'2012*, Tours.

Bouissou M., Bon J.L., 2003, A new formalism that combines advantages of fault-trees and Markov models: Boolean logic Driven Markov processes. *Reliability Engineering and System Safety*, 82 (2), 149-163.

Bouissou M., 2006, Détermination efficace de scenarii minimaux de défaillance pour des systèmes séquentiels, *15ème Congrès de Maîtrise des Risques et de Sûreté de Fonctionnement, Lambda-Mu'2005*, Lille.

Coudert O., Madre J.C., 1992, Implicit and incremental computation of primes and essential primes of Boolean functions, *Proceedings of the 29th ACM/IEEE Design Automation Conference, DAC'92.*

Dutuit Y., Rauzy A., 2005, Approximate estimation of system reliability via fault trees. *Reliability Engineering and System Safety*, 87 (2), 163-172.

Carton O., 2008, *Langages formels, calculabilité et complexité.* Ed. Vuibert.

Cassandras C.G., Lafortune S., 2008, *Introduction to Discrete Event System, Second Edition.* Springer.

Chaux P.Y, Roussel J.M., Lesage J.J., Deleuze G., Bouissou M., 2013, Towards an unified definition of Minimal Cut Sequences, *4th IFAC Workshop on Dependable Control of Discrete Systems (DCDS 2013)*, York (United Kingdom).

Garg V.K., Kumar R., Marcus S.I., 1999, A Probabilistic Language Formalism for Stochastic Discrete-Event Systems. *IEEE Trans. on Automatic Control,* 44(2), 280-293.

Ibanez-Llano C., Rauzy A., Melendez E., Nieto F., 2010, A reduction approach to improve the quantification of linked fault trees through binary decision diagrams. *Reliability Engineering and System Safety*, 95 (12), 1314-1323.

Kumar R., Garg V. K., "Control of Stochastic Discrete Event Systems Modeled by Probabilistic Languages", *IEEE Trans. on Automatic Control*, vol. 46, no. 4, 2001, 593-606.

Papazoglou I.A., 1998, Mathematical foundations of event trees. *Reliability Engineering and System Safety,* 61 (3), 169-183.

Perez Castaneda G. A., Aubry J.F., Brinzei N., 2011, Stochastic hybrid automata model for dynamic reliability assessment. *Proceedings of the Institution of Mechanical Engineers Part O Journal of Risk and Reliability*, 225 (1), 28-41.

Rauzy A., 1993, New algorithms for fault tree analysis, *Reliability Engineering and System Safety,* 40, 203-211.

Wang X., Ray A. 2004, A language measure for performance evaluation of discrete event supervisory control systems. *Applied Mathematical Modelling* no. 28: 817-833.