Measurement of Fairness in Process Models Using Entropy and Stochastic Petri Nets

Martin Ibl
Institute of System Engineering and Informatics, University of Pardubice, Pardubice, Czech Republic

Keywords: Fairness, Stochastic Petri Nets, Entropy, Behavioural Analysis, Process Measure.

Abstract: Measurements of various properties of the process models in the last few years become relatively widely explored area. These are properties such as uncertainty, complexity, readability or cohesion of process models. Quantification of these properties can provide better insight in term of, for instance, user-friendliness, predictability, clarity, etc. of the process model. The aim of this work is to design a method for quantification of fairness in the process models which are modelled using stochastic Petri nets. The method is based on mapping the set of all reachable markings of Petri net into Markov chain and then quantification of entropy from stationary probabilities of the individual places (all places or a specific subset). The resulting value of fairness is from the interval <0, 1>.

1 INTRODUCTION

There are a number of modelling languages that are used to describe business processes. Individual modelling languages differ from each other mainly in notation, modelling ability (power), mathematical foundation, etc. In recent years a several metrics began to disseminate that measure specific characteristics in the models, which were developed by specific modelling languages. These properties are for instance uncertainty (Jung et al., 2011, Ibl, 2013), complexity (Lassen and van der Aalst, 2009) or cohesion (Reijers and Vanderfeesten, 2004). The measurement of these properties can be used for evaluation of variety assumptions, which are relative to the model and outline more detailed information about their structure and behaviour. The evaluation of these properties can provide useful information during analysis of, for instance, user-friendliness, understandability, usability, maintainability and other (González et al., 2010).

In this work is proposed a procedure for the analysis and evaluation of fairness in the models created using stochastic Petri nets. Quantification of fairness in process models can implies a number of factors such as the overload of nodes, bottleneck, starvation, etc. Increasing of fairness in the process model can lead to more effective static and dynamic characteristics of the modelled system / process.

Petri nets (Petri, 1962) are an appropriate tool for modelling processes, which are characterized by a non-determinism, synchronization, parallelism and concurrency. Stochastic Petri nets (Molloy, 1982) extends the classic Petri nets with the possibility to allocate exponential distribution to each transition, which allows to refine the behaviour of the modelled system and also to execute various performance analysis.

According to ("Fairness", 2011), the concept of fairness is defined as: „The quality of treating people equally or in a way that is right or reasonable".

When analysing reactive and concurrent systems (Völzer and Varacca, 2012), this term is understood more generally and formally, for instance, fairness does not need to be in connection with people, but any abstract entities (machinery, departments, communication channels, etc.). The term "treating" specifies any process, which is in association with these entities. From this perspective, it is then possible to divide these entities into those for which is the process defined (e.g., packets in the process of network communication) and those, which provide the process (e.g., switches and communication channels in the process of network communication). On the fairness can thus be regarded as a prerequisite (predicate), which is related to the system / process. The classic view of fairness in reactive and concurrent systems is associated with a number of (usually infinite) occurrences of some

Ibl M..
Measurement of Fairness in Process Models Using Entropy and Stochastic Petri Nets.
DOI: 10.5220/0005098301150120
In *Proceedings of the 9th International Conference on Software Paradigm Trends* (ICSOFT-PT-2014), pages 115-120
ISBN: 978-989-758-037-6

specific events (e.g., thread A is waiting to enter the critical section) in an infinite sequence of events (execution), which can occur in a given system (Baier and Katoen, 2008).

One of the most widespread theories of fairness is linked to the issue of formal specification and verification. In this area, the concept of fairness is regarded as a temporal property, which, along with safety and liveness represents the behavioural properties that are verified on the system / model (e. g., model checking) (Baier and Katoen, 2008).

Another area, which deals with the notion of fairness, is the queuing theory (Larson, 1987, Palm, 1953). In this area was always preferred the examination of performance characteristics such as throughput, response time, queue length, etc. In the last decade, however, the fairness has begun to be more important. According to (Raz et al., 2004), the fairness in the queuing theory is divided into fairness with regard to the order (time of arrival), time of service and operational deployment.

The above described “classic” view of farness is associated with the entities whose progress (concurrency) is modelled (e.g., processes in an operating system, customers in a supermarket, users of a website, material in the manufacture, etc.). The second way to look at fairness is associated with the entities that provide the process (e.g., machinery, employees, processors, servers, etc.). In this work presented view of fairness is associated with this second way. Fairness in this context is understood as the uniformity of the workload of specific entities that provide the process. Fairness as defined in the area of scheduling, formal verification or Petri nets is usually a categorization. This means that the assumption is verified and the result is only to determine whether a system / model meets this assumption or not (True or False). The approach presented in this work allows quantifying fairness in terms of uniformity of workload, i.e. what specific proportion of time each entity has against other entities (portion of time each state of system occurs against all states.

The aim of this work is to define a method that allows quantifying the fairness of stochastic Petri net models. This objective is achieved using the concepts of information theory (Shannon´s entropy (Shannon, 1948)) and stochastic processes (Markov chains).

2 STOCHATIC PETRI NETS

The following is the definition of stochastic Petri nets (Molloy, 1982) and few basic concepts that will be needed in the following . A solid introduction to stochastic Petri nets can be found in (Marsan, 1990).

Definition 2.1: Stochastic Petri net is a 5-tuple, $SPN = (P, T, F, \Lambda, W, C, M_0)$ where:

- $P = \{p_1, p_2, p_{3,}, \dots, p_m\}$ – a finite set of places,
- $T = \{t_1, t_2, t_3, \dots, t_n\}$ – a finite set of transitions,
- $P \cap T = \emptyset$ – places and transitions are mutually disjoint sets,
- $F \subseteq (P \times T) \cup (T \times P)$ – a set of edges, defined as a subset of the set of all possible connections,
- $\Lambda: T \rightarrow R^+$ – an exponentially distributed firing rate of transitions,
- $W: F \rightarrow N_1$ – a weight function, defines the multiplicity of edges,
- C:P → N_1 – capacities of places,,
- $M_0: P \rightarrow N_0$ – an initial marking.

Definition 2.2: Marking of Stochastic Petri net

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net. Map $M: P \rightarrow N_0$, is called marking of Petri net SPN.

Marking represents the state of the network after execution a specific number of steps, i.e. the firing a specific number of enabled transitions. If a transition is enabled (or not) depends on the net structure and the actual marking.

Definition 2.3: Pre-set, Post-set

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net. Pre-sets and post-sets are defined as:

- ${}^\bullet p = \{t | (t, p) \in F\}$ – the set of input transitions of p,
- ${}^\bullet t = \{p | (p, t) \in F\}$ – the set of input places of t,
- $p^\bullet = \{t | (p, t) \in F\}$ – the set of output transitions of p,
- $t^\bullet = \{p | (t, p) \in F\}$ – the set of output places of t.

Definition 2.4: Enabled transition

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net. Transition $t \in T$ is called enabled with marking M (M-enabled), if

$$\forall p \in {}^\bullet t: M(p) \geq W(p, t)$$

$$\forall p \in t^\bullet: M(p) \leq C(p) - W(t, p)$$

Definition 2.5: Next marking

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net and M is its marking. If a transition $t \in T$ is enabled at marking M, then by its execution is obtained next marking M', which is defined as follows:

$$\forall p \in P: M'(p) = \begin{cases} M(p) - W(p,t), & if\ p \in {}^{\bullet}t \setminus t^{\bullet} \\ M(p) + W(t,p), & if\ p \in t^{\bullet} \setminus {}^{\bullet}t \\ M(p) - W(p,t) + W(t,p), & if\ p \in t^{\bullet} \cap {}^{\bullet}t \\ M(p) & otherwise \end{cases}$$

The situation that the transition t changes the marking M to M', is usually expressed as $M[t\rangle M'$.

Definition 2.6: Sequence of transitions, reachability

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net. Sequence of transitions σ is the sequence of enabled transition that lead from marking M to another marking M'. This situation is denoted as $M[\sigma\rangle M'$. A marking for which there is a sequence of transitions from the initial marking is called reachable marking.

Definition 2.7: The set of all reachable marking

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net and M is its marking. The set of all possible markings reachable from initial marking M_0 in a Petri net SPN is denoted by $R(PN, M_0)$ or simply $R(M_0)$.

$$R(M_0) = \begin{bmatrix} M_0(p_1) & M_1(p_1) & \cdots & M_{|R(M_0|}(p_1) \\ M_0(p_2) & M_1(p_2) & \cdots & M_{|R(M_0|}(p_2) \\ \vdots & \vdots & \ddots & \vdots \\ M_0(p_m) & M_1(p_m) & \cdots & M_{|R(M_0|}(p_m) \end{bmatrix}$$

Definition 2.8: Boundedness

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net. Place $p \in P$ is called k-bounded if:

$$\exists k \in N_1 : \forall M \in R(M_0) : M(p) \leq k$$

Place $p \in P$ is called bounded, if it is k-bounded for some $k \in N_1$. If every place in PN is bounded, then this net is called bounded Petri net.

Definition 2.9: Live marking, live net

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net. Marking $M \in R(M_0)$ is live, if $\forall t \in T$ exist some marking $M_1 \in R(M_0)$ such that transition t is M_1-enabled. If $\forall M \in R(M_0)$ is live, then SPN is live.

3 PROBABILITY OF MARKINGS AND MARKOV CHAINS

The set of all reachable markings can be expressed in terms of Markov chains (Molloy, 1982). For the purposes of defining the stationary probability of each marking $M \in R(M_0)$ is important to define the transition rate matrix.

Definition 3.1: Transition rate matrix

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a Petri net and $R(M_0)$ its reachability set. Transition matrix **Q** of Petri net SPN is defined as:

$$\boldsymbol{Q}: \big(R(M_0) \times R(M_0)\big) \rightarrow \mathrm{R}$$

Where values are made according following rule and the matrix A form right stochastic matrix:

$$\boldsymbol{Q}_{i,j} = \begin{cases} \displaystyle\sum_{t_k \in \{h : h \in T \wedge M_i \geq 1 \wedge M_i[h\rangle M_j\}} \lambda_k, & if\ i \neq j \\ \displaystyle -\sum_{k=1}^{|R(M_0)|} Q_{i,k}, & if\ i = j \end{cases}$$

Definition 3.2: Stationary probabilities

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a Petri net and $\boldsymbol{Q}$ is its transition rate matrix. Stationary distribution vector $\boldsymbol{\eta}$ is defined as normalized left null space of transition matrix $\boldsymbol{Q}$:

$$\boldsymbol{\eta Q} = \boldsymbol{0}$$
$$\boldsymbol{\eta 1}^T = 1$$

Vector $\boldsymbol{\eta}$ then represents the probability of each SPN marking:

$$\boldsymbol{\eta} = \begin{bmatrix} \Pr(M_0) \\ \Pr(M_1) \\ \vdots \\ \Pr(M_{|R(M_0)|}) \end{bmatrix}$$

Definition 3.3: Long term probability of marking $M \in R(M_0)$ is defined as a corresponding element of vector $\boldsymbol{\eta}$:

$$\eta_i = \Pr(M_i)$$

The probability of marking M can be seen as a joint probability of markings of individual places:

$$Pr(M) = \Pr(M(p_1) = x_1, M(p_2) = x_2, \dots, M(p_m) = x_m)$$

When calculating the stationary probabilities it is appropriate to check whether the model fulfil the liveness property, since each dead marking of Petri net corresponds to absorb state in terms of Markov chains. Each absorption state can always occur, i.e. its probability equal 1 and thus all live markings have probability equal 0. This would lead to a fully deterministic model without any uncertainty.

4 ENTROPY AND FAIRNESS

Entropy (Shannon, 1948) can measure the amount of disorder, which is associated with a random variable.

Definition 4.1: Entropy

$$H(S) = -\sum_{i=1}^{n} Pr(s_i)\log_2 Pr(s_i)$$

$$0 \cdot \log_2(0) \equiv 0$$

where S represents the system (random variable) and s_i its states (the specific values of a random variable).

Definition 4.2: Maximum entropy

The maximum entropy of the system S, which can be in n different states can be expressed as follows:

$$H_{\max}(S) = \log_2 n$$

Definition 4.3: Stationary probability of places

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net and $\boldsymbol{\eta}$ is the vector of the stationary probabilities (of all reachable markings). Stationary probability vector of individual places $\boldsymbol{\rho^s}$ is defined as:

$$\boldsymbol{\rho^s} = norm(\boldsymbol{M} * \boldsymbol{\eta}) = \frac{\boldsymbol{M} * \boldsymbol{\eta}}{\sum(\boldsymbol{M} * \boldsymbol{\eta})}$$

$$\boldsymbol{\rho^s} = \begin{bmatrix} \Pr(p_0) \\ \Pr(p_1) \\ \vdots \\ \Pr(p_n) \end{bmatrix}$$

where $\boldsymbol{M}$ represents a matrix of all reachable markings. In the case that the specific place in a specific marking comprising more than one token, than the partial probability is multiplied by the number of tokens in that place.

Definition 4.4: Entropy of places

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net and $\boldsymbol{\rho^s}$ its stationary probability vector of places. Entropy of places in SPN is defined as:

$$H_P(SPN) = -\sum_{i=1}^{|P|} {\rho^s}_i \log_2 {\rho^s}_i$$

The same relationship can be used to quantify the entropy of certain subset of places. Let $D^s \subset P$, then the entropy of this subset is defined as:

$$H_P(D^s) = -\sum_{i=1}^{|D^s|} {\rho^s}_i \log_2 {\rho^s}_i$$

Definition 4.5: Fairness of stochastic Petri net

Let $SPN = (P, T, F, \Lambda, W, C, M_0)$ is a stochastic Petri net and $H_P(SPN)$ entropy of its places. Fairness of SPN is defined as:

$$Fairness(SPN) = \frac{H_P(SPN)}{H_{P_{max}}(SPN)}$$

and similarly for $D^s \subset P$:

$$Fairness(D^s) = \frac{H_P(D^s)}{H_{P_{max}}(D^s)}$$

The value of fairness is then located in the interval <0,1>, where 0 represents a fully unfair situation and 1 absolutely fair situation. Higher value fairness indicates the higher uniformity of stationary probability of places and vice versa.

5 EXAMPLE OF SIMPLE MODEL

As a simple example, consider a stochastic Petri net, which is composed of 5 places and 5 transitions, see Figure 1. The model contains some typical elements that are abundant in classic process models. These elements are for instance sequence (transition T_2), AND-split (transition T_1), XOR (transition T_4 and T_5) and cycle (transition T_5). For more information on the mapping of these (and other) elements into Petri net can be found in (Jung et al., 2011).

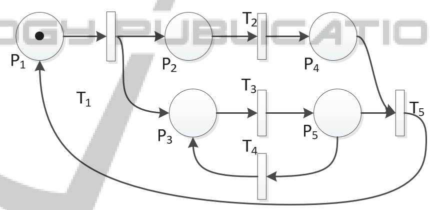

Figure 1: Stochastic Petri net example.

The set of all reachable markings $R(M_0)$ of the Petri net contains five markings:

	M_0	M_1	M_2	M_3	M_4
p_1	1	0	0	0	0
p_2	0	1	0	1	0
p_3	0	1	1	0	0
p_4	0	0	1	0	1
p_5	0	0	0	1	1

When considering the exponential distribution $\Lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)$, the corresponding state space graph (Markov chain) is shown in Figure 2.

This state space corresponds to Markov chain, which generates for $\Lambda = (28, 5, 1, 42, 142)$ the following transition rate matrix:

	M_0	M_1	M_2	M_3	M_4
M_0	-28	28	0	0	0
M_1	0	-47	5	42	0
M_2	0	1	-43	0	42
M_3	0	0	0	-5	5
M_4	142	0	0	1	-143

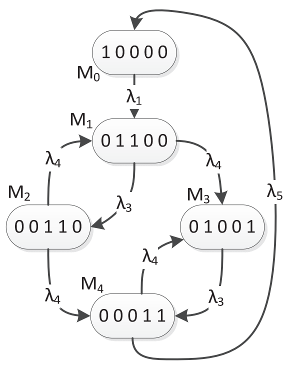

Figure 2: State space / Markov Chain.

The solution of this matrix is a vector of stationary probabilities:

$$\boldsymbol{\eta} = \begin{bmatrix} 0.145 \\ 0.086 \\ 0.010 \\ 0.731 \\ 0.029 \end{bmatrix}$$

The stationary probability of places can therefore be quantified as follows:

$$\Pr(P) = norm(\boldsymbol{M} * \boldsymbol{\eta}) =$$

$$= norm\left(\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} * \begin{bmatrix} 0.145 \\ 0.086 \\ 0.010 \\ 0.731 \\ 0.029 \end{bmatrix}\right) =$$

$$= norm\left(\begin{bmatrix} 0.144 \\ 0.817 \\ 0.096 \\ 0.38 \\ 0.759 \end{bmatrix}\right) = \begin{bmatrix} 0.078 \\ 0.440 \\ 0.052 \\ 0.021 \\ 0.409 \end{bmatrix}$$

It is then possible to quantify the fairness of the presented example:

$$Fairness(SPN) = \frac{H_P(SPN)}{H_{P_{max}}(SPN)} =$$

$$= -(0.078 \log_2 0.078 + 0.440 \log_2 0.440 +$$
$$+0.052 \log_2 0.052 + 0.021 \log_2 0.021 +$$
$$+0.409 \log_2 0.409) / \log_2 5 =$$
$$= \frac{1.673}{2.3219} = 0.721$$

This result can be loosely interpreted as the fact that fairness of sample stochastic Petri net reaches 72.1%, which can be classified as a higher degree of fairness. The resulting value is associated with the distribution of the stationary probability of places $\Pr(P)$ – the places P_2 and P_5 have a significantly higher stationary probability than other places.

6 DISCUSSION

Fairness is an important property that is placed on the modelled system. Verification of fairness is associated with a number of specific disciplines such as queuing (Avi-itzhak et al., 2008, Raz et al., 2004), distributed programming (Alpern and Schneider, 1985, Apt et al., 1988) or networking (flow fairness) (Kelly, 1997, Jaffe, 1981). In the most cases, the verification of the fairness is associated with the choice of a suitable predicate whose validity is verified on the model. Another way to verify the fairness is its measurement (preferably as the value from the fixed interval, e.g. (Raz et al., 2004)). These approaches are based on the assessment of fairness for entities that are subjects in the process (users, documents, packets, etc.). In Petri nets are these entities represented as tokens (which are located in different states – places). This paper presents an approach for measuring fairness in stochastic Petri nets using the Shannon entropy and Markov chains. This approach presents the measurement of fairness for entities that provide a specific state of token (place). The value of fairness is influenced by a number of different factors that are associated with network structure (number and distribution of elements such as OR, XOR, AND and LOOP), the values of Λ and initial marking (number and distribution of tokens).

Advantages of this Approach

- The universal method for evaluation of fairness in concurrent and reactive systems. The only requirement is the need for modelling the system using stochastic Petri nets.
- Flexible measuring of fairness among different subsets of places (from one particular place to the whole network).
- The possibility of simulation of the network with various parameters (e.g., number of tokens in an initial marking, firing rate, etc.) a monitor the progress of fairness, i.e. the possibility of finding local optima around a specific operation point (initial values of model parameters).

Disadvantages of this Approach

- Basic shortcomings of Petri nets in general, i.e. state space explosion, restrictions based on the definition, etc.

- The need to specify amongst which places the fairness will be measured and what the resulting value of fairness represents.

7 CONCLUSION AND FUTURE WORK

Measuring of fairness in the process model can be a good indicator for overload detection of different nodes in the model (bottleneck, overwork, etc.)
In this paper has been defined method for calculating the fairness of any process model, which can be modelled by stochastic Petri net. Defined method can be applied to a specific sublet of places, as well as the whole Petri net.
The actual fairness quantification is based on the measurement of entropy from steady-state probabilities of all places (or a specific subset of places). On the prime example is presented the calculation of the fairness.

The future research will be focused on defining this method using coloured Petri nets, which allow diversification of tokens. This allows measurement of fairness for entities that provide process (states) as well as entities that are subjects in the process (tokens).

ACKNOWLEDGEMENTS

This work was supported by the project No. CZ.1.07/2.2.00/28.032 Innovation and support of doctoral study program (INDOP), financed from EU and Czech Republic funds.

REFERENCES

"Fairness", 2011. *Oxford Advanced American Dictionary for learners of English*
Alpern, B., Schneider, F. B., 1985. Defining liveness. *Information Processing Letters,* 21**,** 181-185.
Apt, K., Francez, N., Katz, S., 1988. Appraising fairness in languages for distributed programming. *Distributed Computing,* 2**,** 226-241.
Avi-Itzhak, B., Levy, H., Raz, D., 2008. Quantifying fairness in queuing systems: Principles, approaches, and applicability. *Probab. Eng. Inf. Sci.,* 22**,** 495-517.
Baier, C., Katoen, J.-P., 2008. *Principles of model checking,* Cambridge, Mass., The MIT Press.
González, L. S., Rubio, F. G., González, F. R., Velthuis, M. P., 2010. Measurement in business processes: a systematic review. *Business Process Management Journal,* 16**,** 114 - 134.
Ibl, M., 2013. Uncertainty Measure of Process Models using Entropy and Petri Nets. ICSOFT 2013 - Proceedings of the 8th International Joint Conference on Software Technologies, 2013. Setubal: Institute for Systems and Technologies of Information, Control and Communication (INSTICC), 542-547.
Jaffe, J., 1981. Bottleneck Flow Control. *Communications, IEEE Transactions on,* 29**,** 954-962.
Jung, J.-Y., Chin, C.-H., Cardoso, J., 2011. An entropy-based uncertainty measure of process models. *Information Processing Letters,* 111**,** 135-141.
Kelly, F., 1997. Charging and rate control for elastic traffic. *European Transactions on Telecommunications,* 8**,** 33-37.
Larson, R. C., 1987. PERSPECTIVES ON QUEUES - SOCIAL-JUSTICE AND THE PSYCHOLOGY OF QUEUING. *Operations Research,* 35**,** 895-905.
Lassen, K. B., Van Der Aalst, W. M. P., 2009. Complexity metrics for Workflow nets. *Information and Software Technology,* 51**,** 610-626.
Marsan, M. A., 1990. Stochastic Petri nets: an elementary introduction. *In:* GRZEGORZ, R. (ed.) *Advances in Petri nets 1989.* Springer-Verlag New York, Inc.
Molloy, M. K., 1982. Performance Analysis Using Stochastic Petri Nets. *Computers, IEEE Transactions on,* C-31**,** 913-917.
Palm, C., 1953. Methods of Judging the Annoyance Caused By Congestions. *TELE (English ed.)***,** 1-20.
Petri, C. A., 1962. *Kommunikation mit Automaten.* Bonn: Institut für Instrumentelle Mathematik, Schriften des IIM Nr. 2.
Raz, D., Levy, H., Avi-Itzhak, B., 2004. A resource-allocation queueing fairness measure. *SIGMETRICS Perform. Eval. Rev.,* 32**,** 130-141.
Reijers, H. A., Vanderfeesten, I. T. P., 2004. Cohesion and Coupling Metrics for Workflow Process Design. *In:* DESEL, J., PERNICI, B. & WESKE, M. (eds.) *Business Process Management.* Springer Berlin Heidelberg.
Shannon, C. E., 1948. A mathematical theory of communication. *Bell System Technical Journal,* 27**,** 379-423.
Völzer, H., Varacca, D., 2012. Defining Fairness in Reactive and Concurrent Systems. *Journal of the ACM,* 59**,** 13-13:37.