확률 추상 시간 기계

박지연°, 노경주, 이문근
전북대학교 컴퓨터과학과
e-mail : {jypark, kjuno, mklee}@cs.chonbuk.ac.kr

Probabilistic Abstract Timed Machine

Ji-yeon Park°, Kyoung-ju Noh, Moon-kun Lee
Dept. of Computer Science, Chonbuk National University

요 약

ATM(*Abstract Timed Machine*)은 실시간 시스템을 순환공학에서 명세하기 위해 고안된 LTS(*Labeled Transition System*)이다. ATM은 상태 기반 명세 언어의 문제점이라 할 수 있는 상태폭발을 해결하기 위해 새로운 개념의 모드를 정의하였으며, 아키텍처 기반으로 구성되어 시스템이 이해를 높일 수 있도록 하였다. 또한 통신, 시간, 예외처리 등의 다양한 실시간 시스템의 속성을 표현한다. 그러나 실시간 시스템은 이 외에도 자원의 한계와 통신 수행의 경쟁 등으로 인해 많은 제약을 가진다. PATM(*Probabilistic ATM*)은 자원의 사용이나 통신의 수행에 많은 경쟁을 하는 시스템의 동작의 성공과 실패를 확률로써 표현하여 시스템의 동작을 예측하며 내고장성(*fault tolerance*)을 동적으로 수행할 수 있도록 하기 위해 ATM을 확률의 개념으로 확장한 것이다. 본 논문에서 내고장성은 환경 요소를 파라미터로하는 확률 함수에 의해 이루어 진다. 본 논문에서는 기존 ATM을 간단히 소개하고, PATM이 필요한 이유와 PATM의 정의, PATM이 동적인 상황에서 어떻게 작용하는지에 대해 기술한다.

1. 서론

일반적으로 실시간 시스템은 규모면에서 방대할 뿐만 아니라 시간, 예외처리 및 통신 등의 다양한 속성을 가지고 있다. ATM[1]은 방대한 규모의 실시간 시스템을 효율적으로 명세, 이해, 검증하고자 고안된 정형 기법이다. 본 논문에서는 ATM을 내고장성을 위해 확장한 정형 기법에 대해 기술한다.

많은 실시간 시스템들은 보다 효율적이고 정확하며 빠른 활용을 위해 분산환경에서 동작한다. 분산환경에서 동작하는 시스템을 개발하고, 배치하기위해서는 많은 비용이 소요되며 시스템을 개발할 때 최소의 비용으로 최대의 효과를 목적을 달성하고자 한다. 즉, 시스템의 개발과 운용에 대한 비용과 자원이 제한되어 있으며 제한된 환경에서 최적의 시스템을 명세하고, 분석할 수 있어야 한다. 자원과 자본의 제약은 시스템 실행시 뜻하지 않은 고장을 유발할 수 있다.

제한된 환경에서 실행될 수 있는 동작의 확률을 제공하고 분석함으로써 내고장성과 동적인 시스템의 동작 분석을 수행하기 위해 확률 ATM, 즉, PATM을 고안하였다.

PATM(*Probabilistic ATM*)은 자원의 사용이나 통신의 수행에 많은 경쟁을 하는 시스템의 동작의 성공과 실패를 확률로써 표현하여 시스템의 동작을 예측하며 내고장성(*fault tolerance*)을 동적으로 수행할 수 있도록 하기 위해 ATM을 확률의 개념으로 확장한 것이다. 본 논문에서 내고장성은 환경 요소를 파라미터로하는 확률 함수에 의해 이루어 진다. 본 논문에서는 기존 ATM을 간단히 소개하고, PATM이 필요한 이유와 PATM의 정의, PATM이 동적인 상황에서 어떻게 작용하는지에 대해 기술한다.

본 논문은 2절에서는 비교연구, 3절에서는 기존의 ATM에 대해 소개한다. 4절에서는 PATM을 정의하고, PATM의 주요 특징인 확률의 동적 수정에 대해 기술하고 이에 대한 분석의 예를 살펴본다. 마지막으로 5절에서는 결론 및 향후 연구를 기술한다.

본 연구는 한국과학재단 특정기초연구(과제번호 1999-2-203-003-3) 지원으로 수행되었음.

2. 비교연구

확률을 명세 단계에서 표현 가능하도록 하는 명세 기법으로는 *Timed Automata*를 확률의 개념으로 확장한 PTA(*Timed Automata with Probability*), 실시간 프로세스 알지브라인 ACSR[4]을 확장한 PACSR[3], CCS[6]를 확장한 PCCS[5]와 실시간 시스템을 위한 TCCS[5]와 조합하여 확장한 PTCCS[5] 등이 존재한다. 이 명세 기법들은 확률에 대한 명세기법과 확률과 관련된 각종 명세 속성들을 정의, 기술한다.

그러나, 확률에 대한 명세 기법을 제공하는 대부분의 정형 기법들은 확률을 기술하고 분석할 때 정적이 방법을 사용한다. 즉, 모든 환경변수를 고려한 확률 값을 명세 전에 파악한 뒤 시스템 명세에 기술한다. 이렇게 기술된 명세로부터 각종 속성을 분석하고 정의한다. 일단 확률이 정해지면 확률 값을 재 수정하지 않는 동안에는 항상 같은 확률에 근거한 분석이 이루어지게 되는 것이다.

많은 시스템들이 분산환경에서 다양한 형태로 통신을 수행하고, 제한된 자원을 사용하는 환경에서 확률은 매 순간 변하게 된다. 따라서, 고정된 확률을 통해서 시스템을 분석하는 것은 실세계의 변화 속성을 만족시키지 못한다.

PATM은 기존 명세 기법들처럼 정적인 확률을 제공할 뿐만 아니라 실행 과정에서 변경 가능한 동적인 확률 표현을 제공한다. 실행 과정에서 동적으로 변화하는 확률은 내고장성에 대한 장점을 제공한다. 확률 초기 명세 값에 근거하여 실행중인 시스템은 고장이 발생했을 때 관련 확률을 변경하여 다시 실행할 경우 고장을 발생시키지 않은 다른 전이를 수행할 수 있다.

3. ATM

ATM은 소프트웨어 재 역공학 과정에서 소프트웨어를 명세하기 위해 고안된 LTS이다. ATM은 임무 위급 시스템과

같은 실시간 시스템을 명세, 분석 및 검증을 수행 할 수 있는 정형 기법이다.

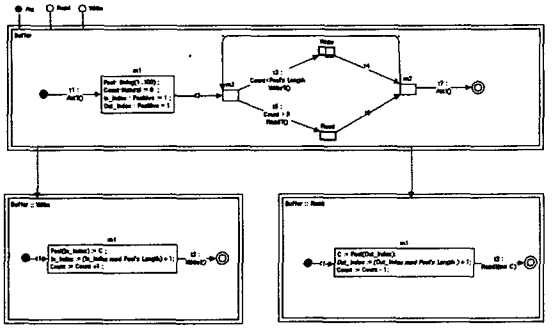
3.1 ATM 구성 요소

ATM은 모드(mode)의 집합, 가드(guard)된 전이(transition)의 집합, 포트(port), 실행 시작점과 실행의 종료점으로 구성된다.

머신은 ATM의 기본 단위로 내부에 모드를 포함한다. 일반적으로 태스크, 프로시저와 같은 독립적 프로그램 블록 단위를 표현 한다.

모드는 시스템의 상태를 나태기 위한 것으로 모드의 역할에 따라 각 모드는 타입화 되어 이름, 아이디, 타입, 시간 제약 등 자신을 대표할 수 있는 속성을 갖는다. 모드의 종류로는 계산 모드, 추상화 모드, 주제 모드가 있다. 계산 모드는 실행문을 내부에 직접적으로 포함하고 있어 자체의 실행에 의해 다음전이를 유발한다. 추상화 모드는 소프트웨어 구조 계층의 하위 단계의 머신을 현재단계에서 추상화 시킨다. 주제 모드는 소프트웨어의 특정 기능을 수행하거나 특정 기능에 영향을 받는 모드들을 표현하는 모드이다.

전이는 이벤트, 조건, 시간제약으로 구성된 레이블을 갖고, 머신 간의 통신을 위해 포트를 가지는데 활성화(activation), 엔트리(entry), 대체(substitute)의 세 종류 포트가 있다. <그림 1>은 ATM의 명세 예제를 보여준다.



<그림 1> ATM 명세 예제 (Buffer)

3.2 ATM 정의

ATM은 정형적으로 $M = <CAS, S, F, T, P>$로 정의 된다. M은 ATM의 기본 단위인 머신을 말하여, CAS는 머신 내 모드의 집합을 말한다. CAS는 계산 모드 C와 추상화 모드 A, 주제 모드 S로 구성된다. $CAS=<N, St, R>$로 구성된다. N은 각 모드의 이름, St는 모드에 포함되는 실행문, R은 시간 제약이다. 시간에 대한 제약 R은 준비 시간, 주기, 실행 시간, 데드라인의 속성을 갖게 된다. S는 머신의 시작점 집합, F는 머신의 종료점의 집합을 말한다. P는 머신과 머신 간의 메시지나 데이터의 송수신이 일어나는 포트로 활성화 포트, 엔트리 포트, 대체 포트로 구성된다. T는 모드와 모드간의 또는 모드와 머신 간의 전이의 집합이다. 전이는 레이블 $L = < Co, E, R>$을 갖는다. 레이블에서 Co는 조건, E는 이벤트, R은 전이의 제약 시간을 나타낸다.

4. PATM(*Probabilistic ATM*)

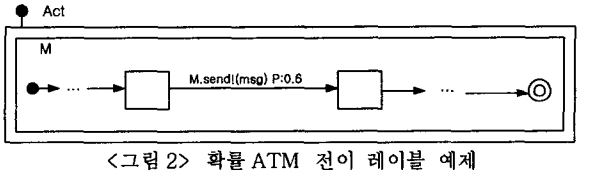
2절 비교연구에서도 언급하였듯이 기존의 연구는 확률의 정적 표현을 통하여 시스템을 분석하고 있다. 시스템이 동작하는 중에 변화하는 확률을 명세에 반영하지 못하며 따라서 항상 변화 가능한 실세계의 시스템을 정확하게 분석할 수 없다는 단점이 있다. PATM은 기존의 확률 명세 기법을 보완하여 (1)고정된 확률 값을 표현하고 추가적으로 전이상에 (2)확률 함수 $f(x_1,...x_n, y_1,...y_m)$를 기술한다.

확률 함수는 확률에 영향을 주는 환경 요인들을 파라미터로 한다. 이러한 환경 요인으로는 통신양(traffic), 머신 간의 연결이 한번에 전달할 수 있는 메시지의 양, 머신 간을 연결해준 매개체의 활성화 여부, 통신을 위한 경로의 수 등이 있다. 다양한 환경 요인은 값이 변하는 것과 변하지 않는 것으로 구분할 수 있으며 변하는 것은 $x_1,...x_n$으로 표현하고 변하지 않는 것은 $y_1,...y_m$으로 한다.

4.1 PATM 정의

확률 ATM은 ATM과 같은 $M = <CAS, S, F, T, P>$로 정의 되며 전이에 기술되는 레이블 형식이 확률을 표현하기 위해 추가정의 된다. 확률 ATM의 전이 $T \subseteq \{CAS \times L \times CAS, M \times M\}$는 (source, label, target)인 형태를 취한다. Label L은 기존 ATM의 정의 $L = < Co, E, R>$에 확률이 추가된 $L = < Co, E, R, P >$로 구성된다. Co는 조건, E는 이벤트, R은 전이의 제약시간을 나타내며 P가 해당 전이가 가지는 전이 확률 값 또는 확률 함수 f를 표현한다. 한 모드에서 발생하는 확률의 합은 1이다.

<그림 2>는 ATM머신 M의 채널 *send*를 통하여 메시지 *msg*를 보내며 이 때 메시지가 제대로 전송될 확률이 0.6임을 보여주는 레이블이다.



<그림 2> 확률 ATM 전이 레이블 예제

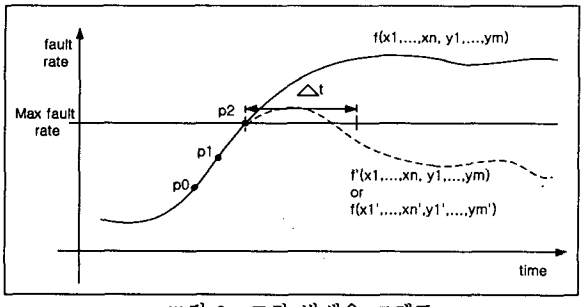
4.2 PATM의 확률 명세와 내고장성

PATM은 확률을 명세할 때 고정된 값을 표현하거나 가변적 함수 값을 제공하는 확률 함수를 사용하여 시스템이 가질 수 있는 확률을 명세 한다. PATM에 고정된 확률 값을 명세하여 분석하는 방법은 다른 확률 정형 기법과 크게 다르지 않으며 고정된 값이 명세된 ATM을 통하여 시스템의 정적 분석을 수행한다. 그러나, 다른 명세 기법과 달리 확률 함수 f를 사용함으로써 PATM은 명세된 시스템의 동적 분석을 제공하며 실행에 대한 확률의 동적 수정을 제공한다. 이는 시스템에 대한 내고장성과 밀접한 관련을 가진다.

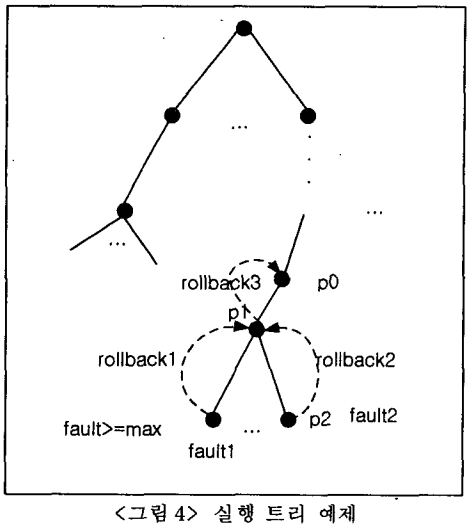
PATM의 확률 함수를 고려했을 때 시스템의 내고장을 위한 대책은 크게 3가지로 나눌 수 있다. (1) 고장이 발생했을 때 이전 상황으로 되돌아가(rollback) 선택 가능한 다른 전이를 수행하는 것, (2) 되돌아가 고장을 발생시킨 확률 함수를 변경함으로써 전이 확률을 변경하여 내고장성을 높이는 것, (3) 고장을 발생시킨 전이의 한단계 이전으로 더 되돌아 새로운 전이를 발생시키는 것이다. (3)의 경우는 고장을 발생시키지 않은 전이를 수정함으로써 시간과 비용면에서 많은 낭비를 초래함으로 (1)과 (2)의 경우를 사용한다.

고장과 고정 확률 값과 확률 함수에 관해 예를 들어 살펴보자. <그림 3>은 시간상에서 변화되는 고장 발생 비율과 최대 고장 허용치를 나타내는 예제 그래프이다. 그림에서

시스템이 허용하는 최대 고장 비율을 *max fault rate*가 나타낸다. 초기 확률 값으로부터 예상되는 고장 발생 비율은 확률 함수 $f(x_1,...x_n, y_1,...y_n)$에 의해 발생한 실선으로 표현된다. 시스템에 대한 요구사항은 동작 중 최대 고장 비율 허용치를 넘지 않아야 하며 만약 최대치를 넘었을 경우 Δt시간 이내에 고장 발생 비율을 낮춰야 한다. 즉, 그래프의 점선 그래프를 유도하는 새로운 확률 함수로 변경해야 한다.



<그림 3> 고장 발생율 그래프



<그림 4> 실행 트리 예제

Case 1: 고장과 고정 확률

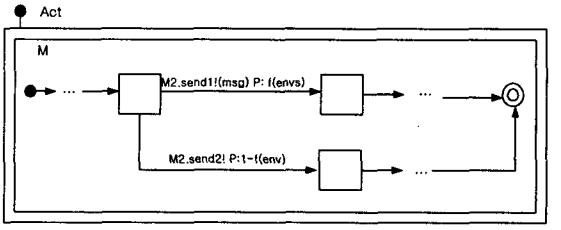
<그림 2>에 기술된 것처럼 확률이 전이 상에 고정된 값으로 표현되어 있는 경우에 발생한 경우를 <그림 4>의 "fault1"이 보여준다. 이때 시스템은 "rollback1"을 수행하여 고장을 발생시킨 전이와는 다른 전이를 선택하여 고장에 대응하게 되며 기존의 명세 기법이 이러한 방법을 취한다.

Case 2: 고장과 확률 함수

확률 함수는 실행 중 발생하는 고장에 대해 동적인 대응을 하기위한 방법을 제공한다. <그림 4>의 실행 트리에서 *p2*는 확률 함수를 갖는 전이에 의해 발행한 고장 지점을 나타낸다. *p2*지점은 <그림 3>에서 *p2*의 지점에 대응된다. 시스템은 "rallback2"를 수행하여 확률 함수를 변경함으로써 확률 값을 조정하여 적절한 내고장성을 수행한다.

이 때, 확률 함수를 조정하는 방법은 두 가지가 있다.

첫째는 확률 함수 $f(x_1,...x_n, y_1,...y_m)$를 고장 발생 비율을 낮추는 $f'(x_1,...x_n, y_1,...y_m)$로 변경하는 방법이고, 둘째방법은 확률 함수의 환경 요소인 파라미터의 값을 변경하거나, 추가, 삭제한 $f(x'_1,...x'_n, y_1,...y_m)$ 으로 수정하여 <그림 3>의 점선 그래프와 갖은 요구사항을 만족하는 함수 값을 얻을 수 있다.



<그림 5> 확률 함수를 전이의 레이블로 갖는 ATM

그러나, 시스템에서 확률에 의한 전이에 의해 발생한 고장이 모두 확률 함수를 변경함으로써 고장을 허용하는 완전한 시스템을 기술하는 것은 아니다.

시스템이 가진 여러 하드웨어나 소프트웨어적 환경 중에는 고장이 발생하면 동작을 멈추고 고장을 해결해야 하는 부분이 있다. 따라서 확률 함수를 변경하고자 할 때나, 확률 함수를 ATM에 기술할 때, 또는 고정된 확률 값을 계산할 때 실제 시스템이 운용되는 환경에서 동적으로 변경될 수 없는 요인, 또는 동적으로 변경되지 않아야 하는 요인들을 찾아내어 이들이 미치는 영향이 사전에 반영되거나, 혹은 고장 발생이후 보완되어야 한다.

5. 결론 및 향후 연구

본 논문에서는 실시간 시스템을 명세하기 위해 고안한 ATM이 실세계에서 동작하는 시스템을 보다 현실적으로 명세하며, 사실적인 분석을 할 수 있도록 하기 위해 확률을 통해 확장한 PATM에 대하여 기술하였다.

PATM을 통하여 자원의 사용과 통신의 수행 과정에서 발생할 수 있는 동작의 성공과 실패의 확률을 명세할 수 있으며 동작 중인 시스템의 최대 고장 허용치에 대한 요구사항을 확률 함수를 통하여 동적으로 변경하도록 하였다.

향후 연구로는 확률에 영향을 주는 환경 요소와 이 요소 중 동적 변경에 영향을 주는 요소와 변경 요소로 작용할 수 없는 요소 등을 파악해야 하며, 확률을 기반으로 실행되는 시스템의 실행 분석을 위한 추가 연구도 수행되어야 한다. 확률을 위한 다양한 정의와 속성 기술, 실행에 관한 분석 및 분석 결과의 효율적 이용에 대해 이루어질 것이다.

참고문헌

[1] 노경주, 박지연, 이문근, "*순환공학을 위한 정형기법 : 추상시간 기계*", 한국정보과학회 소프트웨어공학연구회 2000 정형기법워크샵, pp.137-148.

[2] Anna Philippou, Oleg Sokolsky, Insup Lee, "*Specifying Failures and Revoeries in PACSR*", http://www.cis.upenn.edu.

[3] Hanene Ben Abdallah, "*Grhphical Communicating Shared Resources : a Language for the Specification, Refinement and Analysis of Real-Time Systems*", Ph.D. Thesis, The University of Pennsylvania, 1996.

[4] Hans A. Hansson, "*Time and Probability in Formal Design of Distributed Systems*", ELSEVIER, pp.37~78., 1994.

[5] Robin Miner, "*Communication and Concurrency*", Prentice Hall., 1992.