

Systemy czasu rzeczywistego

Mariusz Żynel

`mariusz@math.uwb.edu.pl`

`http://math.uwb.edu.pl/~mariusz/`

Uniwersytet w Białymstoku

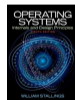
2024/2025

Dlaczego musi być *egzamin*?

- Warunkiem dopuszczenia do egzaminu jest uzyskanie zaliczenia laboratorium na ocenę pozytywną
- Egzamin pisemny składa się z 30 pytań w formie testu
- Do każdego pytania są 3 propozycje odpowiedzi, ale tylko jedna jest poprawna
- Odpowiedź na każde z pytań testowych warta jest 0 lub 1
- Punktacja:
 - ▶ 5.0 : 27 (91%)
 - ▶ 4.5 : 24 (81%)
 - ▶ 4.0 : 21 (71%)
 - ▶ 3.5 : 18 (61%)
 - ▶ 3.0 : 15 (51%)
- Obowiązuje materiał z wykładu: <http://math.uwb.edu.pl/~mariusz/>

- Trochę historii
- Próbujące systemy danych, kontrolery PID
- Kontrolery wysokiego poziomu, hierarchia kontrolerów
- Przetwarzanie sygnałów
- Twarde i miękkie systemy
- Model referencyjny
- Szeregowanie zadań
- Zasoby i kontrola dostępu do zasobów
- Systemy operacyjne

-  Liu J.W.S.
Real-time Systems
Prentice Hall, New York, 2000
-  Buttazzo G.
Hard Real-Time Computing Systems
Springer, 2023
-  Silberschatz A., Galwin P.B., Gagne G.
Podstawy systemów operacyjnych
WNT, Warszawa 2005
-  Stallings W.
Systemy operacyjne: struktura i zasady budowy
Wydawnictwo Naukowe PWN, Warszawa 2006
-  Stallings W.
Operating Systems: Internals and Design Principles
Pearson, 2014



Co to jest *czas*?

Merriam-Webster

Mierzalny okres, w którym działanie, proces lub stan istnieje lub trwa.

Wiktionary

Nieuniknione przechodzenie zdarzeń z przeszłości do teraźniejszości, a następnie do przyszłości.

Słownik Języka Polskiego

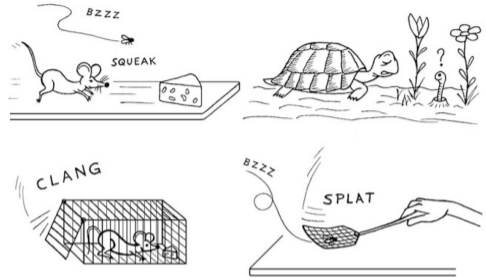
- Nieprzerwany ciąg chwil
- Okres, pora, gdy coś jest wykonywane lub coś się dzieje

1 sekunda

- $1/86400$ średniego dnia w Układzie Słonecznym
- Czas trwania 9192631770 okresów promieniowania odpowiadającego przejściu między dwoma nadsubtelnymi poziomami stanu podstawowego atomu cezu 133

Szybki kontra działający w czasie rzeczywistym

- Żółw jest wolniejszy od myszy w sensie bezwzględnej prędkości, ale zdarzenia, z którymi musi się zmierzyć, są proporcjonalne do działań, które może koordynować, a jest to warunek konieczny, aby każde zwierzę mogło przetrwać w danym środowisku



- Jeśli do środowiska, w którym żyje zwierzę, wprowadzimy zdarzenia rozwijające się szybciej niż zwierzę jest w stanie sobie z tym poradzić, to działania zwierzęcia nie będą już tak skuteczne, a jego przetrwanie zostanie zagrożone
- Pojęcie czasu nie jest nieodłączną właściwością systemu sterowania, ani naturalnego, ani sztucznego, ale jest ściśle związane ze środowiskiem, w którym system ten działa
- Czas należy rozpatrywać w kontekście otoczenia.
- *Szybkie przetwarzanie* – minimalizacja średniego czasu reakcji zestawu zadań
- *Przetwarzanie w czasie rzeczywistym* – spełnienie indywidualnych wymagań czasowych każdego zadania

Otoczenie jest istotnym komponentem każdego systemu

- 4 czerwca 1996, Gujańskie Centrum Kosmiczne
- 12:34:06 UTC, start rakiety Ariane 5 nr 501
- 37s po zainicjowaniu sekwencji startu, na wysokości 3700m, rakieta zbacza o 90° z zaplanowanej ścieżki lotu
- Rakieta doznaje poważnych naprężeń aerodynamicznych, które powodują oderwanie się dodatkowych zbiorników od stopnia głównego
- Wyzwolone zostaje kontrolowane samozniszczenie

...

- Przyczyną awarii był błąd w procedurze inercyjnego systemu nawigacji polegający na przekroczeniu 16-bitowej wartości całkowitej podczas konwersji danych akcelerometrycznych w postaci 64-bitowej wartości zmiennoprzecinkowej
- Błąd konwersji wystąpił dlatego, że oprogramowanie sterujące zostało przeniesione z rakiety Ariane 4, którego dynamika różniła się od dynamiki Ariane 5
- Jeden z najsłynniejszych i najkosztowniejszych błędów oprogramowania w historii wart 370mln USD



Co to jest *pech*?

Czy jesteśmy w stanie go wyeliminować?

Pluskwa w systemie Patriot czy pech?

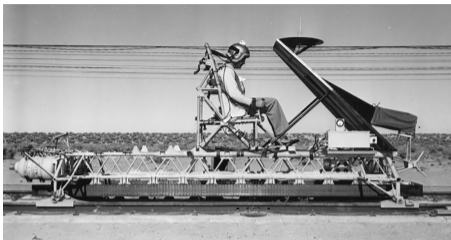
- Gdy radar Patriot wykryje obiekt latający, komputer pokładowy wyznacza jego trajektorię i przeprowadza weryfikację
- Jeśli obiekt przelatuje przez określoną lokalizację, obliczoną na podstawie przewidywanej trajektorii, Patriot zostaje wystrzelony

- 25 lutego 1991, Arabia Saudyjska, Wojna w Zatoce
- Radar Patriot wykrywa pocisk Scud skierowany w Arabię Saudyjską
- Komputer pokładowy wyznacza trajektorię tego pocisku
- Po weryfikacji, zdarzenie klasyfikowane jest jako fałszywy alarm
- Kilka minut później tenże pocisk Scud spada na miasto Dhahran

...

- Z powodu niewielkiego błędu zaokrąglania podczas konwersji czasu z liczby całkowitej na 24-bitową liczbę zmiennoprzecinkową następowało kumulujące się opóźnienie $57\mu\text{s}/\text{min}$
- Ponieważ komputer pracował nieprzerwanie tego dnia od ponad 100 godzin opóźnienie przekroczyło w sumie 343ms
- To opóźnienie spowodowało błąd obliczenia pozycji obiektu o 687m

Jakie prawo stosuje się do opisanej sytuacji?



Projekt MX981 (1949)

Skutki działania na człowieka sił zwalniających o dużej wartości

Prawo Murphy'ego

Jeśli coś może pójść nie tak, to pójdzie nie tak

Stała Murphy'ego

Uszkodzenie przedmiotu jest proporcjonalne do jego wartości

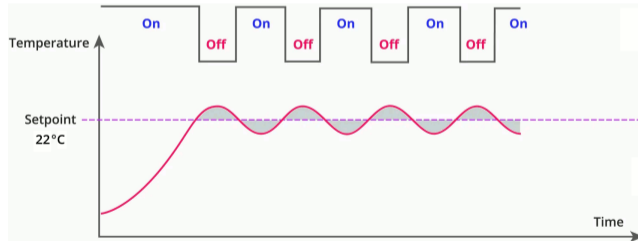
Prawa nie tylko Murphy'ego

- *Prawo Naeser'ego:*
Można stworzyć coś odpornego na bomby, ale nie na pecha
- *Postulaty Troutmana:*
 - ▶ Każdy błąd oprogramowania powoduje maksymalizację szkód
 - ▶ Najpoważniejszy błąd zostanie odkryty 6 miesięcy po testach
- *Prawo Greena:*
Jeśli system jest zaprojektowany tak, by tolerować pewne błędy, to zawsze znajdzie się idiota na tyle utalentowany, aby spowodować nietolerowany błąd
- *Wniosek:* Idioci są bardziej wyrafinowani, niż przedsięwzięte środki, które mają ich uchronić przed szkodą
- *Pierwsze prawo Johnsona:*
Jeżeli system przestaje działać, to zawsze robi to w najgorszym możliwym momencie
- *Drugie prawo Sodda:*
Wcześniej czy później nastąpi najgorszy możliwy zbieg okoliczności
- *Wniosek:* System musi być zaprojektowany tak, aby był odporny na najgorsze możliwe zbiegi okoliczności

Co to jest *regulator PID*?

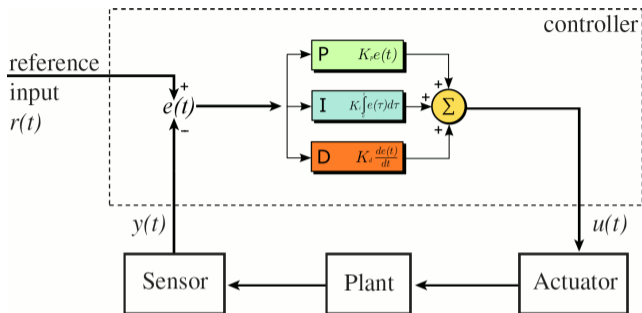
Termostat typu On/Off

- Dom wyposażono w piec centralny z dystrybucją ciepła oraz naścienny, nastawny regulator temperatury (termostat)
- Termostat posiada *sensor* mierzący temperaturę w domu (*process variable PV*) i porównuje ją z ustawianą *wartością zadaną* (*setpoint SP*)
- Gdy temperatura spada poniżej wartości zadanej (*setpoint SP*), piec jest załączany (On)
- Gdy temperatura wzrasta ponad wartość zadaną (*setpoint SP*), piec jest wyłączany (Off)



- Temperatura *oscyluje* wokół ustawionej, co jest nie do przyjęcia w automatyce przemysłowej i sterowaniu ruchem

Regulator PID



- Regulator PID to mechanizm pętli sprzężenia zwrotnego (feedback control loop)
- Oblicza wartość **uchybu** $e(t)$ jako różnicę pomiędzy **wartością zadaną** $r(t)$ (setpoint SP, reference input) i zmierzoną wartością **zmiennej procesu** $y(t)$ (process variable PV, sensor reading):

$$e(t) = r(t) - y(t)$$

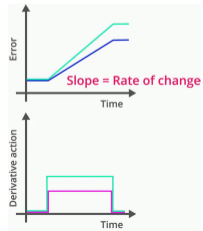
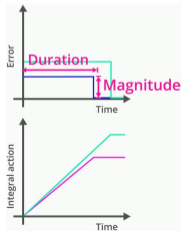
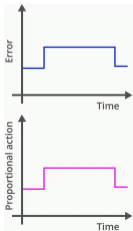
- W celu redukcji uchybu aplikuje, wyliczoną zgodnie z algorytmem PID, **korektę** $u(t)$ na element wykonawczy (actuator)

Algorytm proporcjonalno-całkująco-różniczkujący

- Algorytm obliczania korekty zawiera trzy człony (terms, wyrażenia):
 - proporcjonalny : kompensuje uchyb bieżący
 - całkujący : kompensuje akumulację uchybów z przeszłości
 - różniczkujący : kompensuje przewidywane uchyby w przyszłości
- Stąd określenie PID – Proportional, Integral, Derivative
- Korekta obliczana jest zgodnie ze wzorem:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt},$$

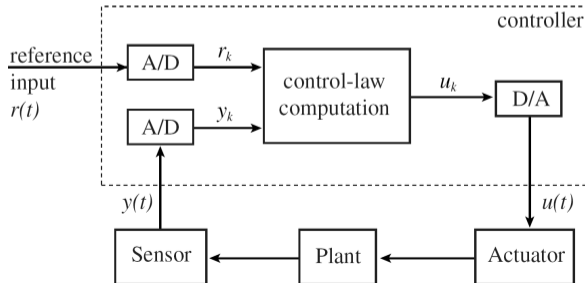
gdzie stałe **współczynniki wzmocnienia** K_p, K_i, K_d dobierane są w drodze **strojenia** regulatora do kontroli danego procesu



Od sterowników analogowych do cyfrowych

	Sterownik analogowy	Sterownik cyfrowy
Czas	ciągły	dyskretny
Stan	ciągły	dyskretny

- Odczyt przez próbkowanie analogowych sensorów (sensor)
- Digitalizacja odczytów (A/D)
- Obliczenie korekty zgodnie z regułami sterowania (control-law)
- Konwersja cyfrowych danych wyjściowych na analogowe (D/A)
- Przesłanie sygnału aktywującego siłowniki (actuator)



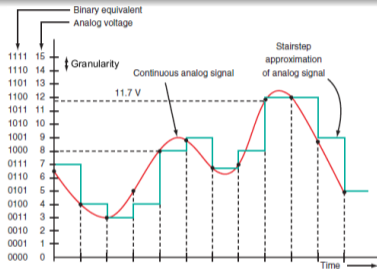
Konwerter analogowo-cyfrowy A/D

Próbkowanie (sampling)

Proces konwersji zmiennego w czasie napięcia na dyskretny w czasie ciąg liczb rzeczywistych

Kwantyzacja (quantization)

Proces mapowania wartości wejściowych z dużego zbioru, zazwyczaj liczb rzeczywistych, na wartości wyjściowe w mniejszym zbiorze, zazwyczaj skończonym zbiorze liczb naturalnych



Konwersja analogowo-cyfrowa (analog-to-digital conversion)

Proces próbkowania i kwantyzacji

Granularność (granularity)

Wielkość odstępu między sąsiednimi wartościami wyjściowymi do wyboru przez kwantyzator

Próbkujące systemy danych

- Regulator cyfrowy operuje na danych dyskretnych
- Okresowo, co T jednostek czasu otrzymujemy zsamplewane wartości

$$r_k, \quad y_k, \quad e_k = r_k - y_k,$$

dla $k = 0, 1, 2, \dots$ będącego kolejnym numerem okresu próbkowania

- W wyniku dyskretyzacji formuły obliczającej $u(t)$ otrzymujemy:

$$u_k = u_{k-2} + \alpha e_k + \beta e_{k-1} + \gamma e_{k-2},$$

gdzie α, β, γ to stałe współczynniki określone w czasie projektowania

- Regulator implementuje się jako nieskończoną pętlę czasową :

```
set timer to interrupt periodically with period T;  
at each timer interrupt, do  
    analog-to-digital conversion to get y;  
    compute control output u;  
    digital-to-analog conversion of u;  
end do;
```

Dobór okresu próbkowania

Okres próbkowania (sampling period)

Czas T pomiędzy dwoma kolejnymi momentami, w których zmienna procesu $y(t)$ oraz wartość zadana $r(t)$ są próbkowane

Częstotliwość próbkowania (sampling rate)

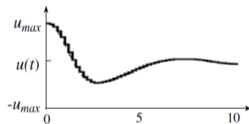
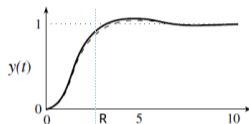
Odwrotność $1/T$ okresu próbkowania wyrażana w Hz

- Okres T jest kluczowym wyborem podczas projektowania
- Zachowanie regulatora istotnie zależy od okresu próbkowania
- Im krótszy okres próbkowania tym bardziej regulator cyfrowy przypomina analogowy, ale zwiększa się zapotrzebowanie na czas procesora
- Przy wyborze okresu próbkowania kluczowe są dwa czynniki:
 - ▶ *Responsywność całego systemu* – chcemy aby urządzenie reagowało bez opóźnień na wydawane polecenia
 - ▶ *Dynamiczne zachowanie urządzenia* – chcemy utrzymać niewielką oscylację w reakcji urządzenia (plant) i zachować kontrolę nad urządzeniem

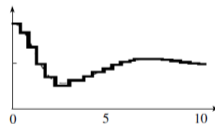
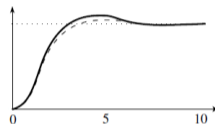
- Często systemem zarządza osoba (np. pilot, kierowca, operator koparki, gracz)
- Operator może wydać polecenie w dowolnym momencie t
- Wynikająca stąd zmiana wartości zadanej (reference input) r odczytywana jest przez sterownik w najbliższym okresie próbkowania, najpóźniej w momencie $t + T$
- Próbkowanie wprowadza opóźnienie w reakcji systemu
- Operator odczuje, że system staje się powolny, gdy opóźnienie przekroczy 1/10 sekundy
- Okres próbkowania danych wprowadzanych ręcznie powinien być poniżej tego limitu

Dynamiczne zachowanie urządzenia

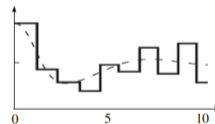
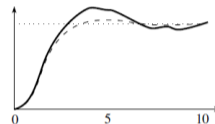
- Urządzenie (plant): ramię głowicy R/W dysku HDD
- Regulator przesuwa ramię do wybranej ścieżki
- W momencie $t = 0$ wartość zadana (reference input) r zmienia się z pozycji początkowej 0 na pozycję docelową 1



$R/T = 20$



$R/T = 8$



$R/T = 4$

- *Linie przerywane* – pozycja $y(t)$ ramienia oraz sygnał sterujący $u(t)$ sterownika analogowego jako funkcje czasu t
- *Linie ciągłe* – pozycja $y(t)$ ramienia oraz analogowy sygnał sterujący $u(t)$ uzyskany z wyjścia cyfrowego regulatora

Czas narastania (rise time)

Czas osiągnięcia przez urządzenie (plant) stanu bliskiego docelowemu w odpowiedzi na zmianę wartości zadanej (reference input) przy zerowych warunkach początkowych


- Im szybciej urządzenie może i musi zareagować na zmiany wartości zadanej (reference input), czyli im krótszy czas narastania, tym szybciej zmienia się sygnał wyjściowy i tym krótszy powinien być okres próbkowania
- Stosunek R/T czasu narastania do okresu próbkowania powinien wynosić od 10 do 20, czyli w czasie narastania powinno mieścić się od 10 do 20 okresów próbkowania
- Krótsze okresy próbkowania mogą mieć niekorzystny wpływ bo błąd *kwantyzacji* rośnie, gdy różnica w kolejnych odczytach próbek analogowych staje się porównywalna lub nawet mniejsza niż *granularność kwantyzacji*

Systemy o wielu częstotliwościach próbkowania (multirate)

- Zazwyczaj stan urządzenia (plant) określony jest nie przez jedną ale *wiele zmiennych*, np. prędkość obrotową, temperaturę, ciśnienie
- Takie urządzenie monitorowane jest przez wiele sensorów i sterowane wieloma elementami wykonawczymi
- Złożony regulator o wielu wejściach i wyjściach to system prostych regulatorów, każdy z jednym wejściem i wyjściem
- Poszczególne zmienne stanu mają zazwyczaj odmienną dynamikę i wymagają różnych okresów próbkowania, np. prędkość obrotowa zmienia się szybciej niż temperatura
- Można do wszystkich regulatorów zastosować najkrótszy okres próbkowania, ale część obliczeń reguł sterowania (control law) będzie niepotrzebnie wykonywana zbyt często
- Powiązane ze sobą logicznie regulatory łączone są w *grupy*
- W jednej grupie okresy próbkowania powiązane są ze sobą *harmonicznie*, tzn. każdy dłuższy okres próbkowania jest całkowitą wielokrotnością każdego krótszego okresu

Struktura sterowania kontrolera lotu helikoptera

Do the following in each 1/180-second cycle:

- Validate sensor data and select data source; in the presence of failures, reconfigure the system
- Do the following 30-Hz avionics tasks, each once every six cycles:
 - ▶ keyboard input and mode selection
 - ▶ data normalization and coordinate transformation, tracking reference update
- Do the following 30-Hz computations, each once every six cycles:
 - ▶ control laws of the outer pitch-control loop 
 - ▶ control laws of the outer roll-control loop
 - ▶ control laws of the outer yaw- and collective-control loop
- Do each of the following 90-Hz computations once every two cycles, using outputs produced by 30-Hz computations and avionics tasks as input:
 - ▶ control laws of the inner pitch-control loop
 - ▶ control laws of the inner roll- and collective-control loop
- Compute the control laws of the inner yaw-control loop, using outputs produced by 90-Hz control law computations as input
- Output commands
- Carry out built-in-test
- Wait until the beginning of the next cycle

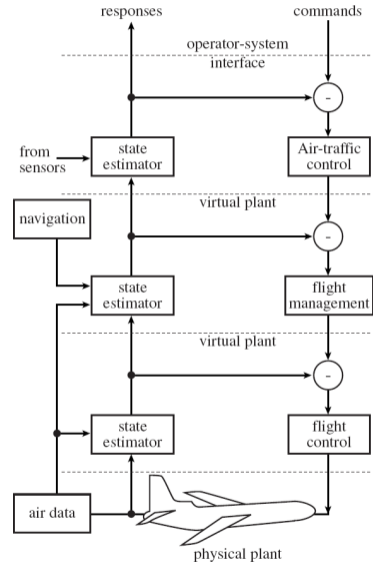
Bardziej złożone obliczenia reguł sterowania

- Prostota regulatora cyfrowego wynika z trzech założeń:
 - ▶ dane z sensorów pozwalają na dokładne oszacowanie wartości zmiennych stanu
 - ▶ dane z czujników określają stan urządzenia (plant)
 - ▶ znane są wszystkie parametry przedstawiające dynamikę urządzenia
- Nie zawsze wszystkie powyższe założenia są spełnione, dlatego regulator powinien być zaimplementowany jako poniższa nieskończona pętla czasowa ∞ :

```
set timer to interrupt periodically with period T;
at each clock interrupt, do
    sample and digitize sensor readings to get measured values;
    compute control output from measured and state-variable values;
    convert control output to analog form;
    estimate and update plant parameters;
    compute and update state variables;
end do;
```

Organizacja regulatorów w złożonych systemach

- W złożonych systemach monitorowania i kontroli regulatory zorganizowane są hierarchicznie
- Przykład: hierarchia kontrolera lotu (flight control), awioniki (flight management) oraz kontroli ruchu powietrznego (air traffic control)
- Kontrola ruchu powietrznego reguluje przepływ lotów do lotnisk docelowych, przypisując każdemu samolotowi czas przylotu do kolejnego punktu orientacyjnego
- Wskazany czas oraz punkt orientacyjny stanowią wartości zadane dla awioniki, która określa ścieżkę lotu tak, aby dotrzeć w odpowiednie miejsce na czas
- Prędkość przelotowa, promień skrętu, szybkość opadania/wznoszenia wymagane do podążania wybraną ścieżką lotu opartą na czasie są wartościami zadanymi dla kontrolera lotu



Sterowanie i kontrola w czasie rzeczywistym

- Regulator na najwyższym poziomie w hierarchii nazywany jest *systemem sterowania i kontroli* (*command and control system*)
- System sterowania i kontroli może mieć niewiele wspólnego z regulatorami niskiego poziomu
- Praca regulatora niskiego poziomu jest czysto okresowa
- System sterowania i kontroli oblicza i komunikuje się w odpowiedzi na sporadyczne zdarzenia i polecenia operatorów
- Wymagania czasowe systemu sterowania i kontroli są mniej rygorystyczne od wymagań regulatora niskiego poziomu
- Regulator niskiego poziomu zwykle działa na jednym lub kilku komputerach połączonych małą siecią lub dedykowanymi łączami
- System sterowania i kontroli jest często dużym rozproszonym systemem zawierającym dziesiątki i setki komputerów oraz wiele różnych rodzajów sieci

Bazy danych czasu rzeczywistego

- Szerokie spektrum systemów informacyjnych: giełda papierów wartościowych, systemy śledzenia i naprowadzania, systemy plików
- Gromadzone dane są *nietrwałe*: bez okresowej aktualizacji jakość danych degraduje się, dane stają się bezużyteczne

Application	Size	Response time		Abs. cons.	Rel. cons.	Permanence
		Average	Maximal			
Air traffic control	20000	0.50 ms	5.00 ms	3.00 sec.	6.00 sec.	12 hours
Aircraft mission	3000	0.05 ms	1.00 ms	0.05 sec.	0.20 sec.	4 hours
Spacecraft control	5000	0.05 ms	1.00 ms	0.20 sec.	1.00 sec.	25 years
Process control		0.80 ms	5.00 sec.	1.00 sec.	2.00 sec.	24 hours

- *Wiek danych* określa jak długo dane nie były aktualizowane
- Zbiór danych jest *bezwzględnie czasowo spójny* (*absolute consistency*), gdy maksymalny wiek danych w tym zbiorze nie przekracza *progu bezwzględnej spójności*
- Zbiór danych jest *relatywnie czasowo spójny* (*relative consistency*), gdy maksymalna różnica wieku danych w tym zbiorze nie przekracza *progu relatywnej spójności*
- Wiek względny bywa bardziej istotny niż wiek bezwzględny: system planowania korelujący natężenie ruchu na autostradzie z natężeniem przepływu pojazdów wjeżdżających i zjeżdżających z autostrady

- Kompresja wideo to proces wymagający dużej mocy obliczeniowej
- W przypadku VOD *kompresja* wykonywana jest *off-line*
- W przypadku telekonferencji kompresja musi być wykonana *on-line*
- *Dekompresja* zawsze wykonywana jest tuż przed prezentacją
- Wymagania czasowe aplikacji multimedialnej wynikają z wymaganej jakości obrazu oraz dźwięku
- Jakość obrazu: *rozdzielczość, frame rate, synchronizacja ruchu ust*
- Im wyższa jakość tym większe zbiory danych, wyższe wymagania na miejsce do przechowywania i przepustowość podczas transmisji oraz dłuższy czas konieczny na kompresję i dekompresję
- Płynny obraz uzyskuje się przy 30 ramkach na sekundę (FPS), przekazy high-definition używają 60 FPS, podczas telekonferencji tolerowalne jest 10-20 FPS
- Aby osiągnąć synchronizację ruchu ust, odstęp między wyświetleniem każdej klatki a prezentacją odpowiadającego jej segmentu audio powinien wynosić nie więcej niż 80 ms i zdecydowanie nie więcej niż 160 ms

Klasyfikacja aplikacji czasu rzeczywistego

Czysto okresowe (regulatory cyfrowe, kontrolery lotu)

- każde zadanie wykonywane jest okresowo
- operacje wejścia/wyjścia przez odpytywanie (polled I/O)
- zapotrzebowanie na zasoby (obliczeniowe, komunikacyjne i pamięci masowej) nie zmienia się znacząco z okresu na okres

Głównie cykliczne (awionika, systemy sterowania procesami)

- większość zadań wykonywana jest okresowo
- system musi reagować asynchronicznie na zdarzenia zewnętrzne

Asynchroniczne i przewidywalne (komunikacja multimedialna, systemy radarowe)

- czas wykonania kolejnych zadań może się znacznie różnić
- duże różnice w zapotrzebowaniu na zasoby w różnych okresach
- zakresy zmian są albo ograniczone, albo znane są statystyki

Asynchroniczne i nieprzewidywalne (inteligentne systemy kontrolne czasu-rzeczywistego)

- zadania o dużej złożoności w czasie wykonywania
- konieczność reagowania na zdarzenia asynchroniczne

> FIN < ACK < FIN > ACK

init 5