Modelamento da Imprecisão Temporal da Observação em Sistemas de Diagnóstico de Ambientes Distribuídos

Volnys Borges Bernal

Tese de doutorado

Depto. de Engenharia de Sistemas Eletrônicos
Escola Politécnica da USP

Orientador: Prof. Dr. Sergio Takeo Kofuji

Agenda

- Motivação e objetivo
- Sistemas de correlação e diagnóstico
- O sistema SMARTS
- Anomalias, sintomas e suas relações
- Modelagem do processo de observação
- Problemas temporais de uma observação
- A nova representação da observação
- Modelagem da observação
- Aglomerado de Intervalos
- Conclusão

Motivação e Objetivo

Motivação

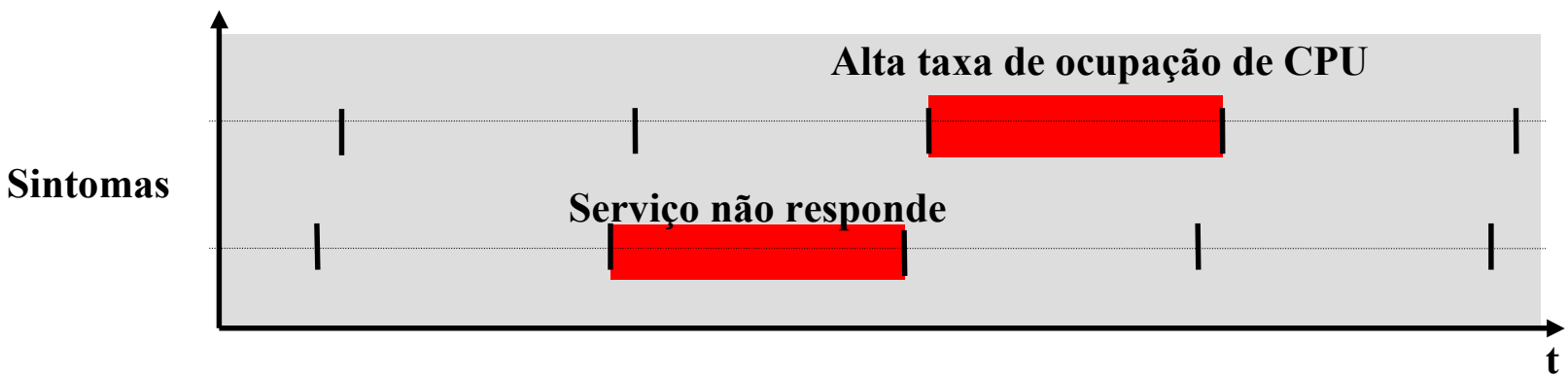
- É impraticável a um operador, mesmo com a utilização de plataforma de gerenciamento, a identificação de problemas em um ambiente distribuído sem um sistema de apoio:
  - Sistema de correlação
  - Sistema de diagnóstico
- Mesmo sistemas de correlação e diagnóstico devem conviver com as seguintes situações:
  - Defasagem temporal da observação
  - Perda de observações
  - Existência de ruídos

Motivação

	Rede de Telecomunicações	Redes de Dados
Protocolo mais utilizado	**CMIP**	**SNMP**
Agente de monitoração	**complexo**	**simples**
Meio de transmissão utilizado no gerenciamento	***In-band*** ***Out-of-band***	***In-band***
Principal método de obtenção de observações	**notificação**	**amostragem periódica (*polling*)**
Observação típica	**não defasada**	**defasada em até 1 ciclo** **defasada em até 2 ciclos**
Perda de observações	**raro**	**freqüente**

Motivação

- Principal problema para detecção de anomalias em redes de dados:
  - **Incerteza temporal da observação**



Objetivo

- Principal
  - Modelagem da imprecisão temporal da observação em sistemas de diagnóstico para sistemas distribuídos e
  - Modelagem da incerteza decorrente da ausência de informações

Sistemas de Correlação de Eventos

Sistemas de correlação de eventos

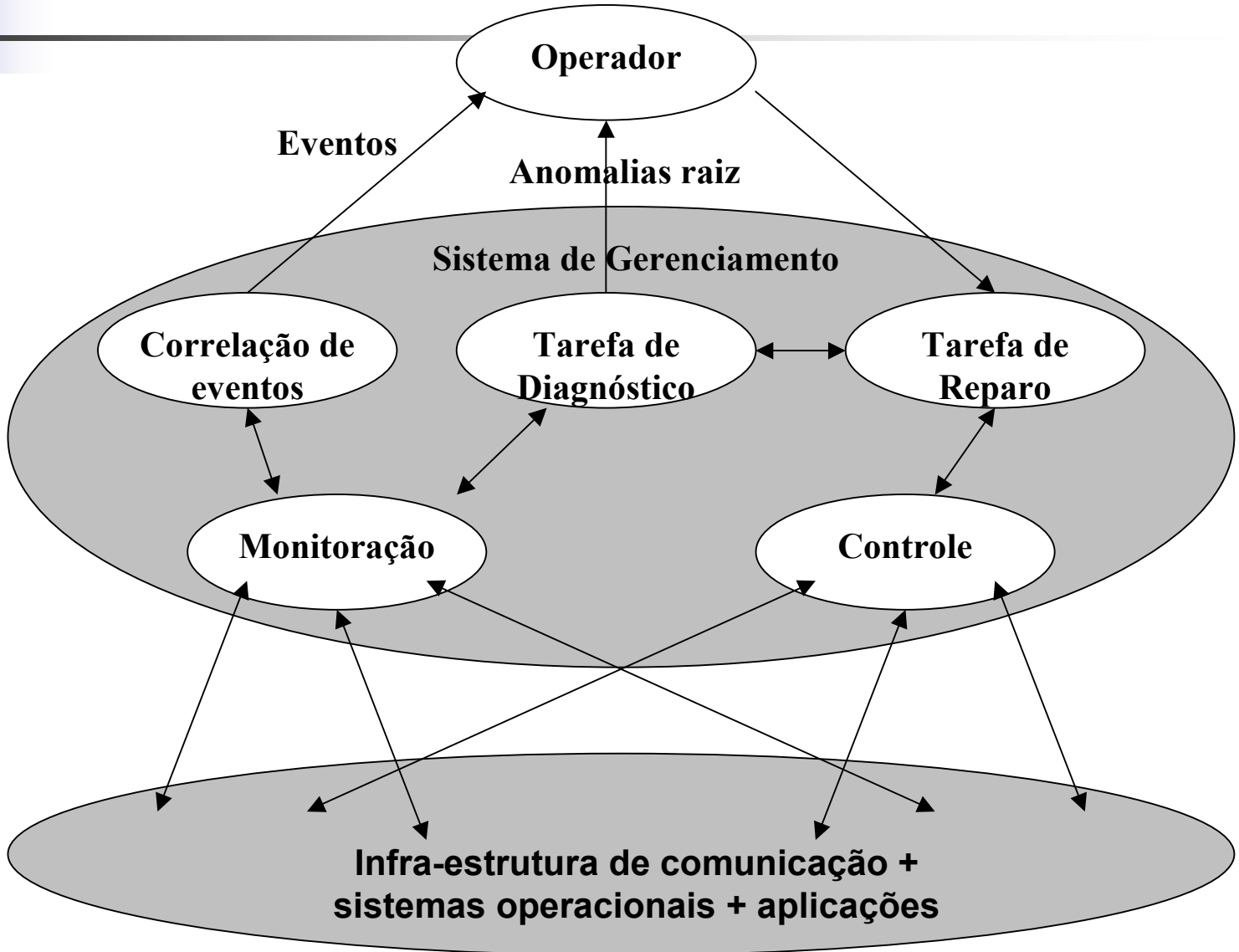
- Correlação de eventos
  - Principal objetivo
    - Reduzir a quantidade de eventos transferidos aos operadores
  - Técnicas utilizadas
    - Compressão, filtragem, supressão, intensificação, …

Sistemas de Diagnóstico

Sistemas de diagnóstico

- Objetivo:
  - Identificar o **conjunto de causas raiz** relacionadas às anomalias que se manifestam através de alguns comportamentos observáveis (sintomas)

Sistemas de apoio

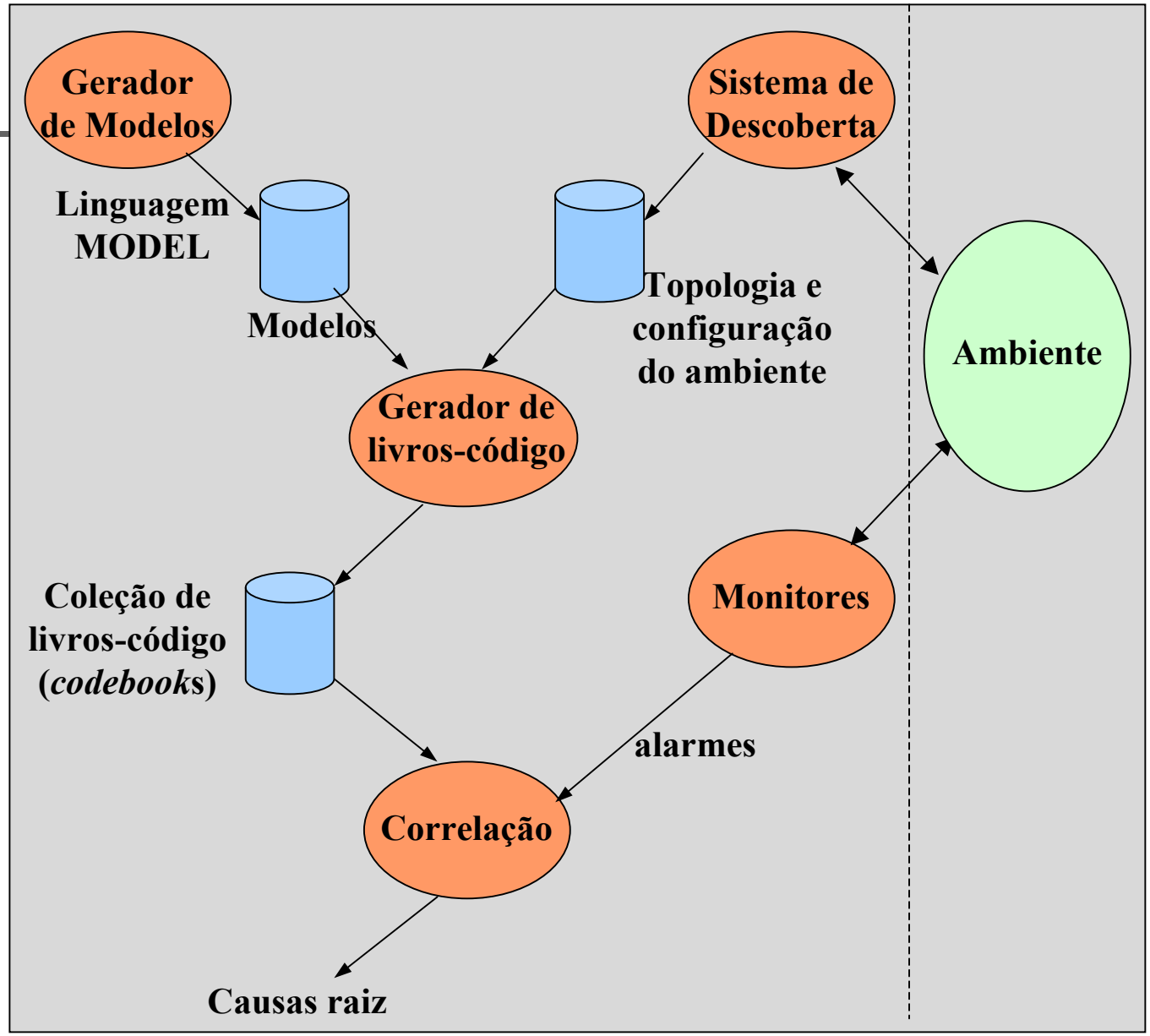


Sistemas de Diagnóstico: Sistema SMARTS

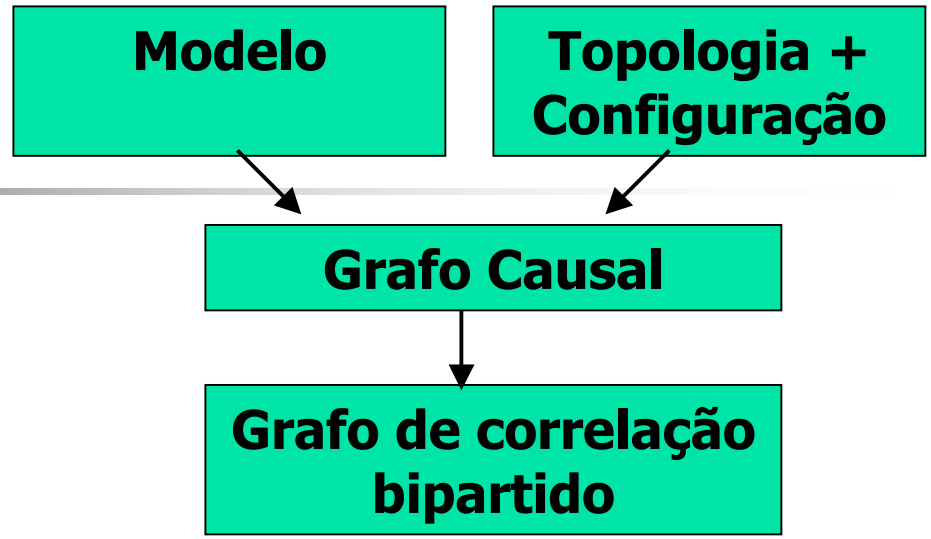
Sistema SMARTS

- Sistema de diagnóstico
- Apoiado por um sistema de correlação
- Utiliza a técnica de correlação por livro-código
- Atemporal
- Baseado em modelo (utiliza a linguagem MODEL para descrição do modelo)

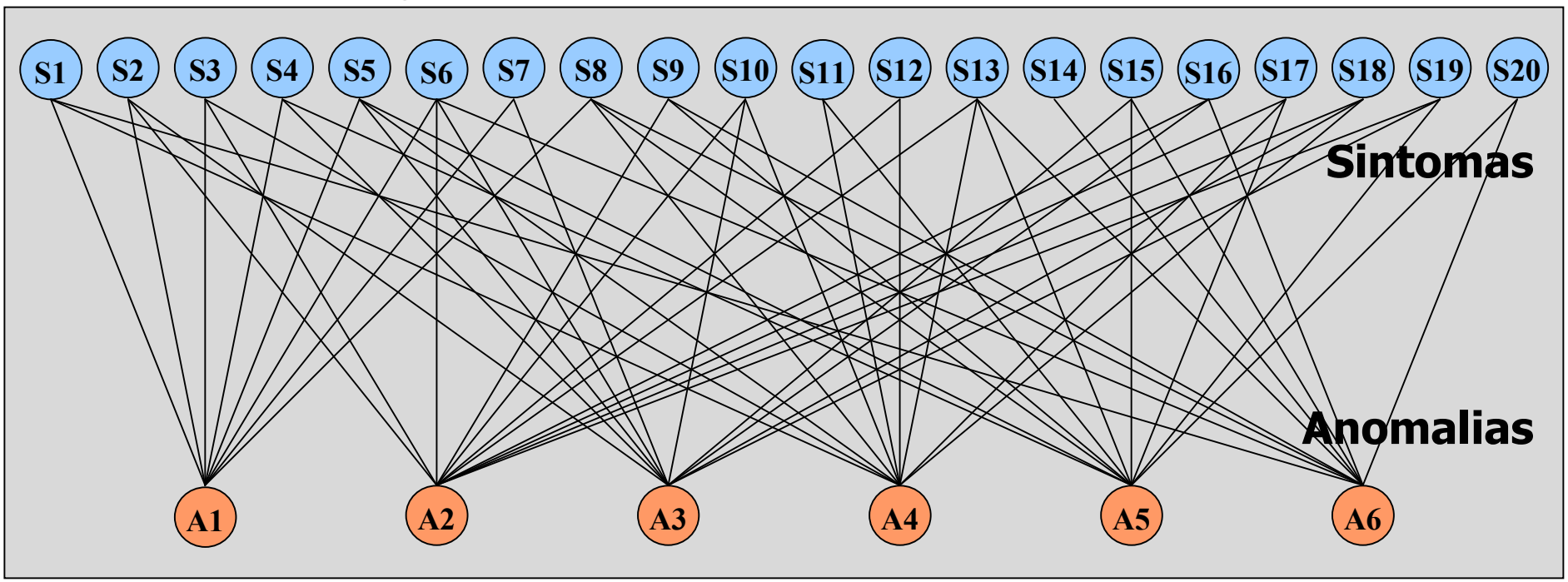
Sistema SMARTS



Sistema Smarts



Grafo de correlação bipartido



Sistema SMARTS

Matriz de correlação

Grafo de correlação bipartido

↓

Matriz de correlação

	Anomalia						ok
Sintoma	A1	A2	A3	A4	A5	A6	
S1	1	0	0	1	0	1	0
S2	1	1	1	1	0	0	0
S3	1	1	0	1	0	0	0
S4	1	0	1	0	1	0	0
S5	1	0	1	1	1	0	0
S6	1	1	1	0	0	1	0
S7	1	0	1	0	0	1	0
S8	1	0	0	1	1	1	0
S9	0	1	0	0	1	1	0
S10	0	1	1	1	0	0	0
S11	0	0	0	1	1	0	0
S12	0	1	0	1	0	0	0
S13	0	1	0	1	1	1	0
S14	0	0	0	0	0	1	0
S15	0	0	1	0	1	1	0
S16	0	1	1	0	0	1	0
S17	0	1	0	1	1	0	0
S18	0	1	1	1	0	0	0
S19	0	1	1	0	1	0	0
S20	0	0	0	0	1	1	0

Sistema SMARTS

Matriz de correlação

↓

Livro-código

	Anomalia						ok
Sintoma	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	
S1	1	0	0	1	0	1	0
S2	1	1	1	1	0	0	0
S4	1	0	1	0	1	0	0

Tolerante à presença de ruído em 1 sintoma →

	Anomalia						ok
Sintoma	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	
S1	1	0	0	1	0	1	0
S3	1	1	0	1	0	0	0
S4	1	0	1	0	1	0	0
S6	1	1	1	0	0	1	0
S9	0	1	0	0	1	1	0
S18	0	1	1	1	0	0	0

Sistema SMARTS

- Principais vantagens
  - Indica com precisão a anomalia raiz (desde que as observações não sejam defasadas em fase e período)
  - Automaticamente computa e atualiza as regras de correlação
  - Extremamente rápido comparado a sistemas baseados em regras
  - Resistente a ruído

Sistema SMARTS

- Principais desvantagens
  - Trata somente falhas únicas
  - Para realizar o correlacionamento todos os sintomas devem estar disponíveis
  - Requer o completo conhecimento das anomalias antes que o livro-código seja computado
  - Não suporta correlacionamento temporal
  - Livro-código deve ser recompilado sempre que o ambiente for alterado
  - **Susceptível a observações defasadas!**

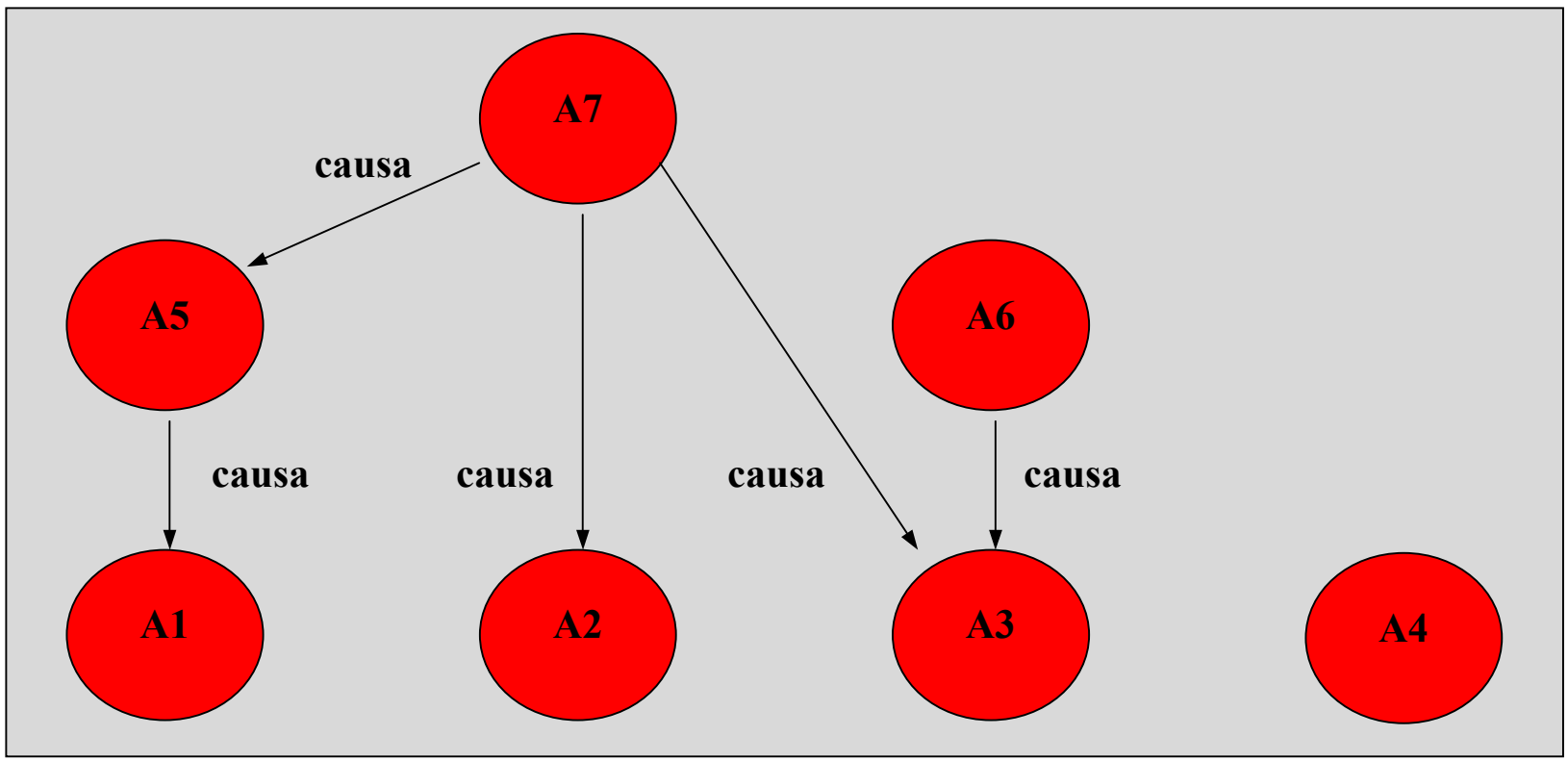
Anomalias, sintomas e suas relações

Anomalias, sintomas e suas relações

- Classificação das relações causais
  - Quanto à possibilidade de causar efeito
    - Necessariamente causa
    - Possivelmente causa
  - Quanto ao retardo do efeito
    - Início imediato
    - Início retardado
    - Término imediato
    - Término retardado
    - Término indeterminado

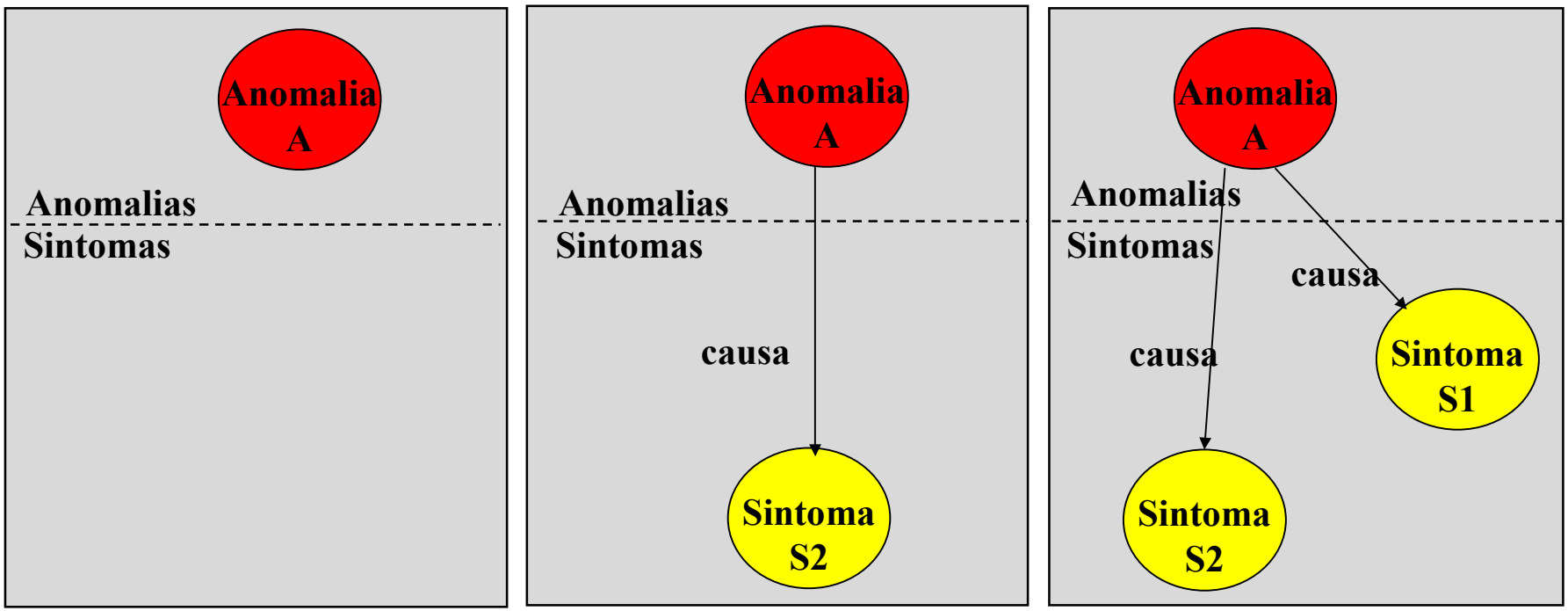
Anomalias, sintomas e suas relações

- Relacionamento causal entre anomalias



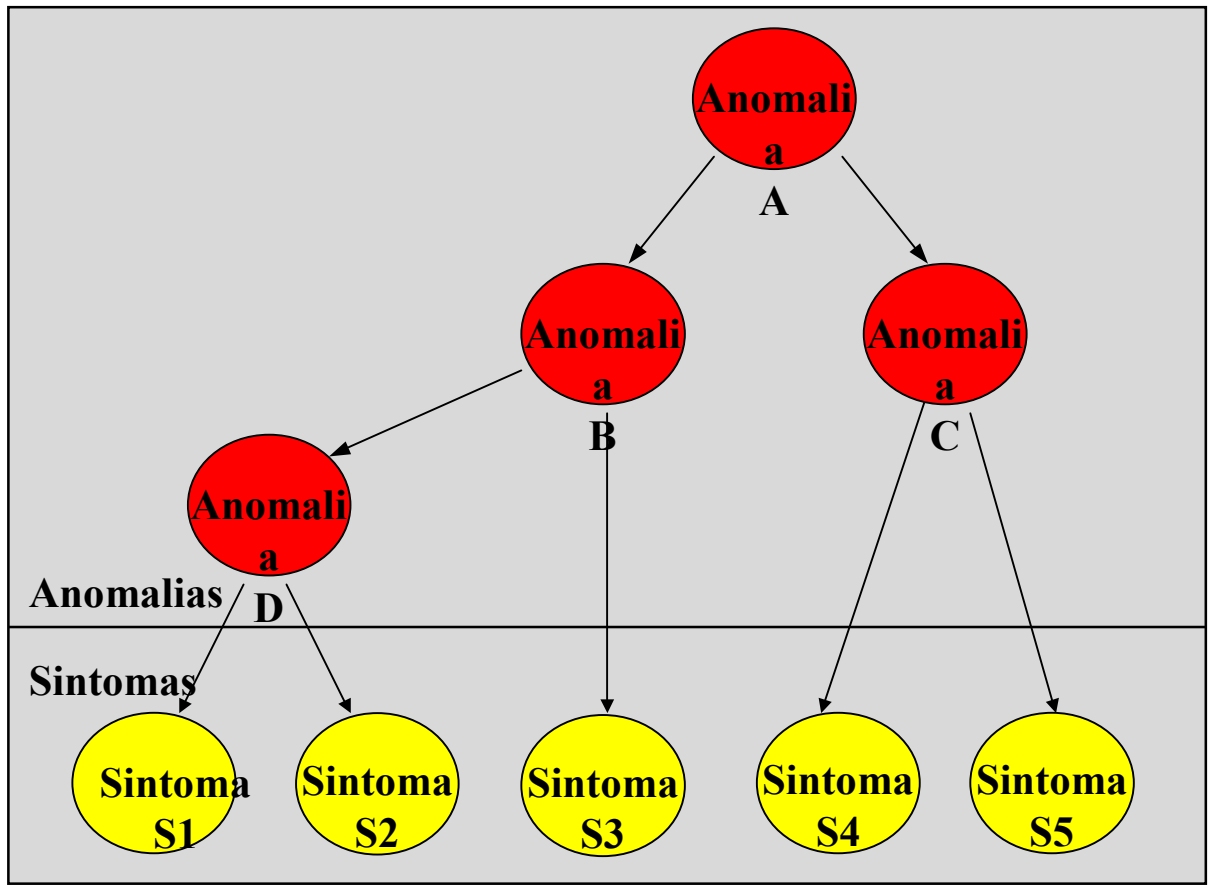
Anomalias, sintomas e suas relações

- Relacionamento causal entre anomalias e sintomas



Anomalias, sintomas e suas relações

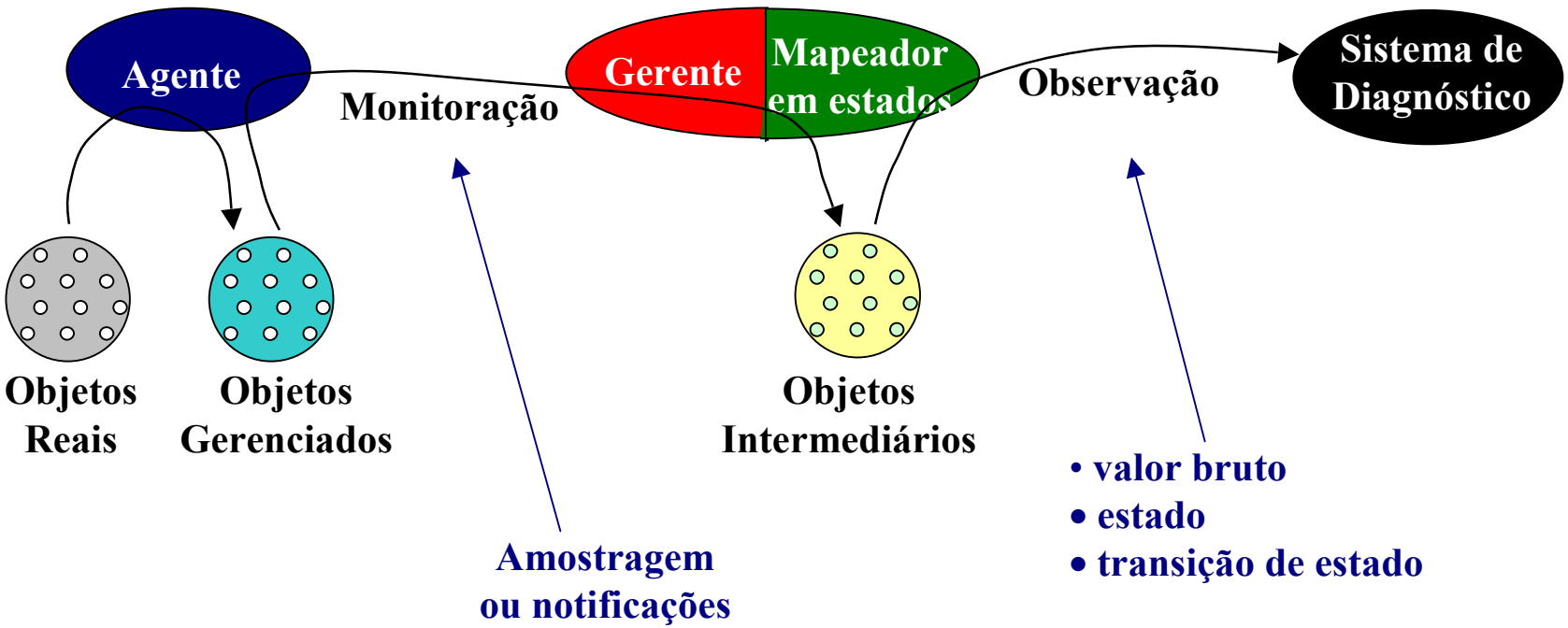
- Relacionamento causal entre anomalias e sintomas



Modelagem do processo de observação

Modelagem do processo de observação

- Visão tradicional do processo de obtenção de observações



Problemas temporais de uma observação

Observação

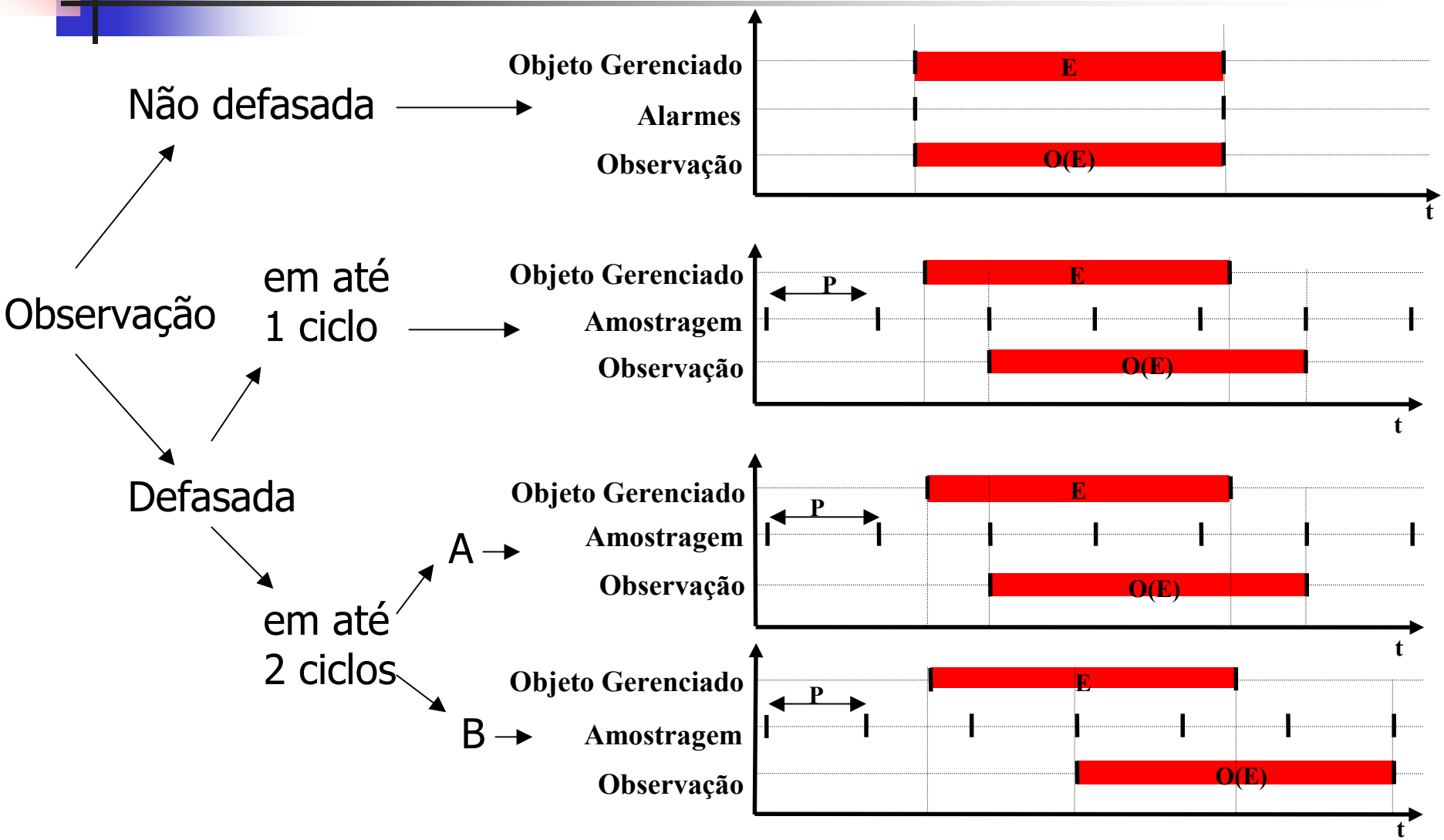
- Problemas temporais

(1) Observação defasada no tempo

(2) Desconhecimento do estado atual de um objeto gerenciado

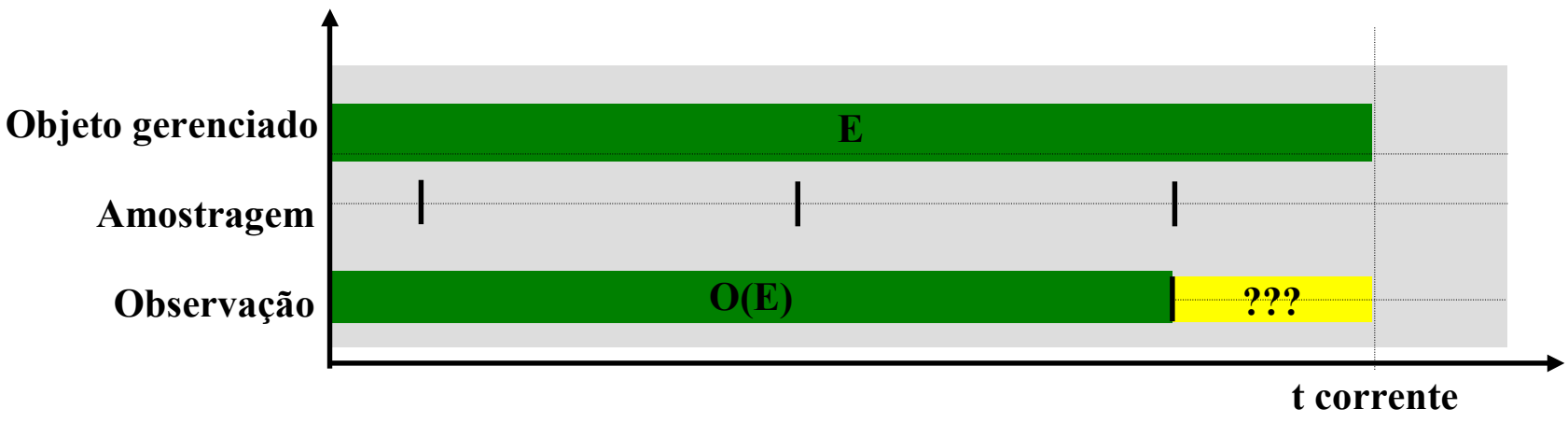
(3) Relacionamento entre observações defasadas no tempo

(1) Observação defasada no tempo



(2) Desconhecimento do estado atual

- Ocorre em:
  - Observações defasadas
- Exemplo:



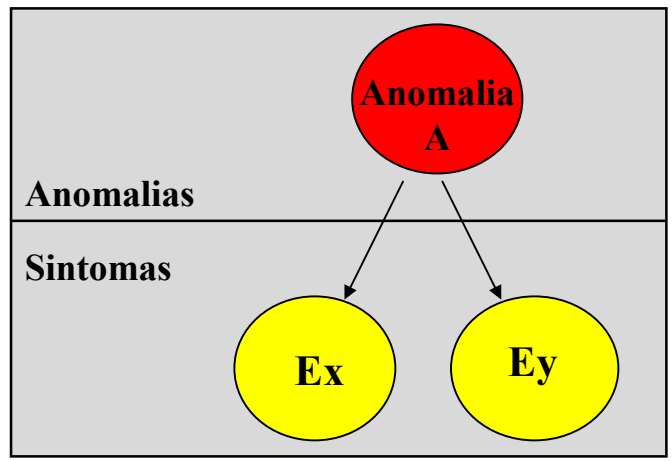
(3) Relac. entre observações defasadas

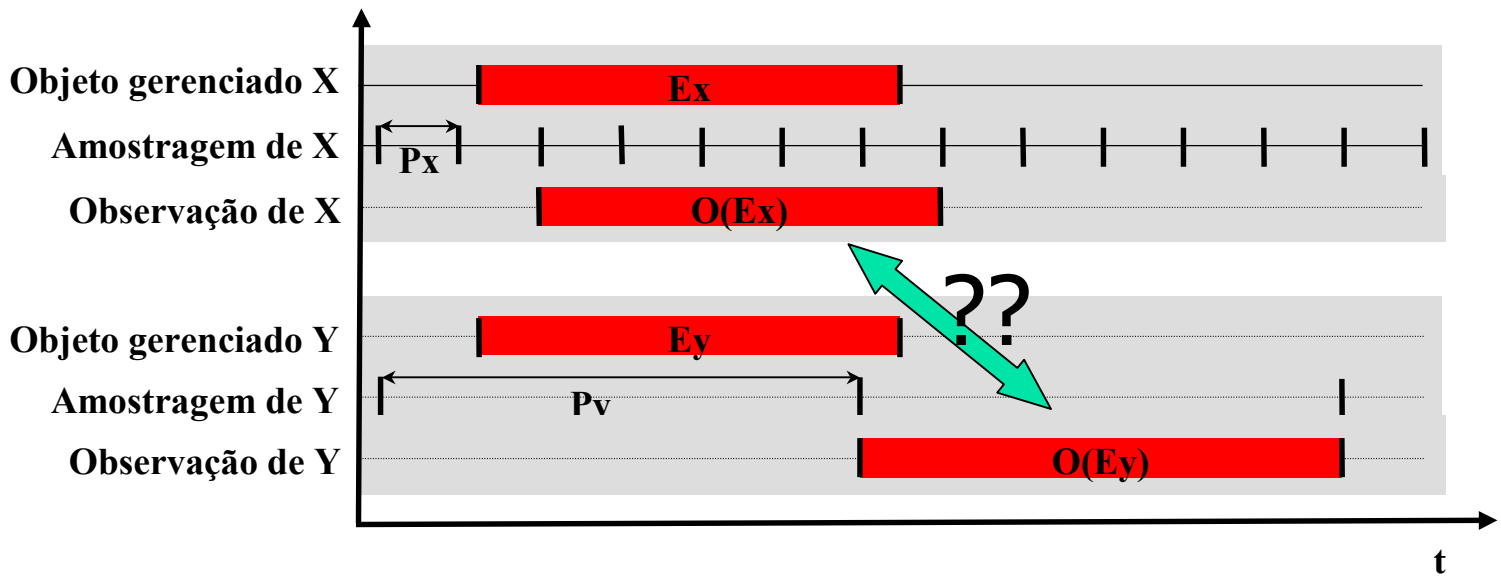
- Não trivial
  - Quando o valor de uma das observações for obtida por amostragem

- Devido a utilização de:
  - Diferentes períodos de amostragem
  - Diferentes fases de de amostragem

(3) Relac. entre observações defasadas

- Exemplo:
  - A causa imediatamente Ex
  - A causa imediatamente Ey
  - O(X) defasada em até 1 ciclo
  - O(Y) defasada em até 1 ciclo

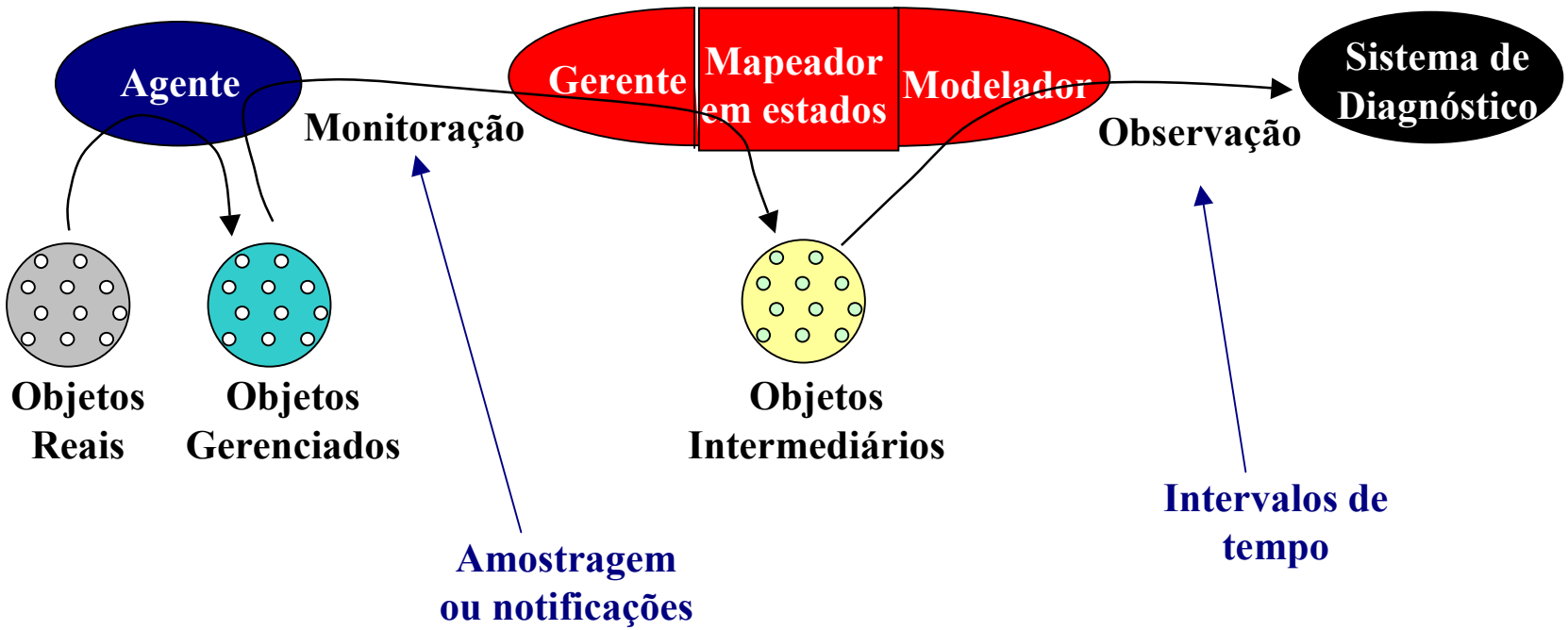




A nova representação da observação

A nova representação da observação

- Visão do novo modelo do processo de obtenção de observações



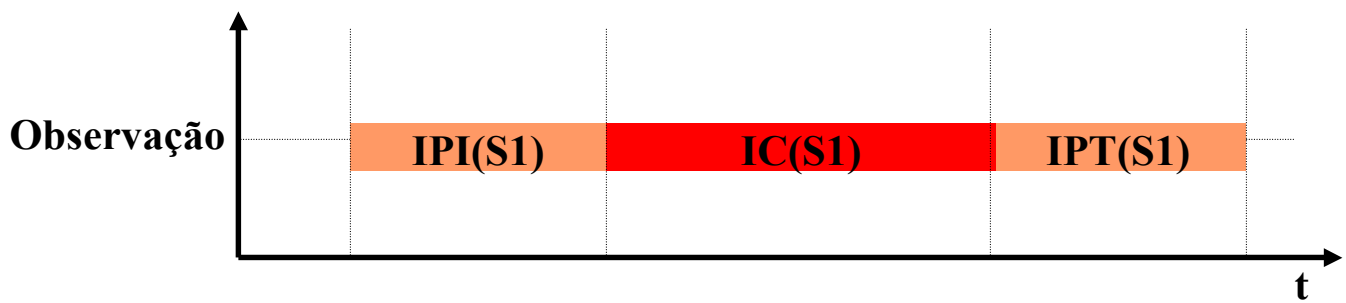
A nova representação da observação

- Objetivo:
  - Representação da localização temporal
  - Explicitar as incertezas temporais

- Observação é composta por
  - Intervalos de possibilidade e certeza
    - intervalo contendo o momento provável da transição de estado
    - Intervalo de certeza da ocorrência do estado
  - Intervalos de incerteza
    - período com desconhecimento do estado
    - intervalos com ausência de monitoração

A nova representação da observação

- Intervalos definidos
  - IC – intervalo de Certeza
  - IPI – Intervalo de Possibilidade de Início
  - IPT – Intervalo de Possibilidade de Término
  - II – Intervalo de Incerteza

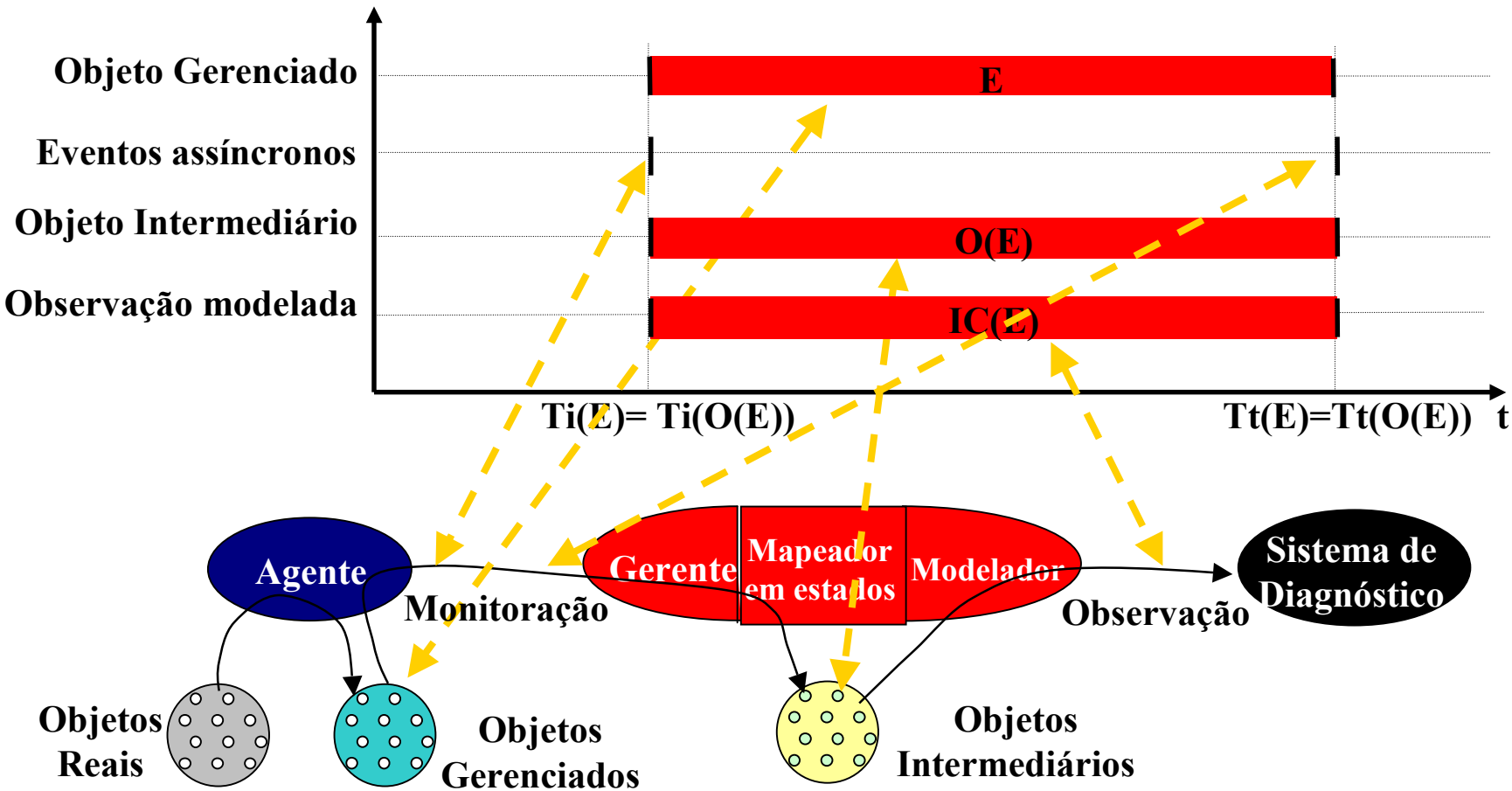


Modelagem da observação

Intervalos de Possibilidade e Certeza

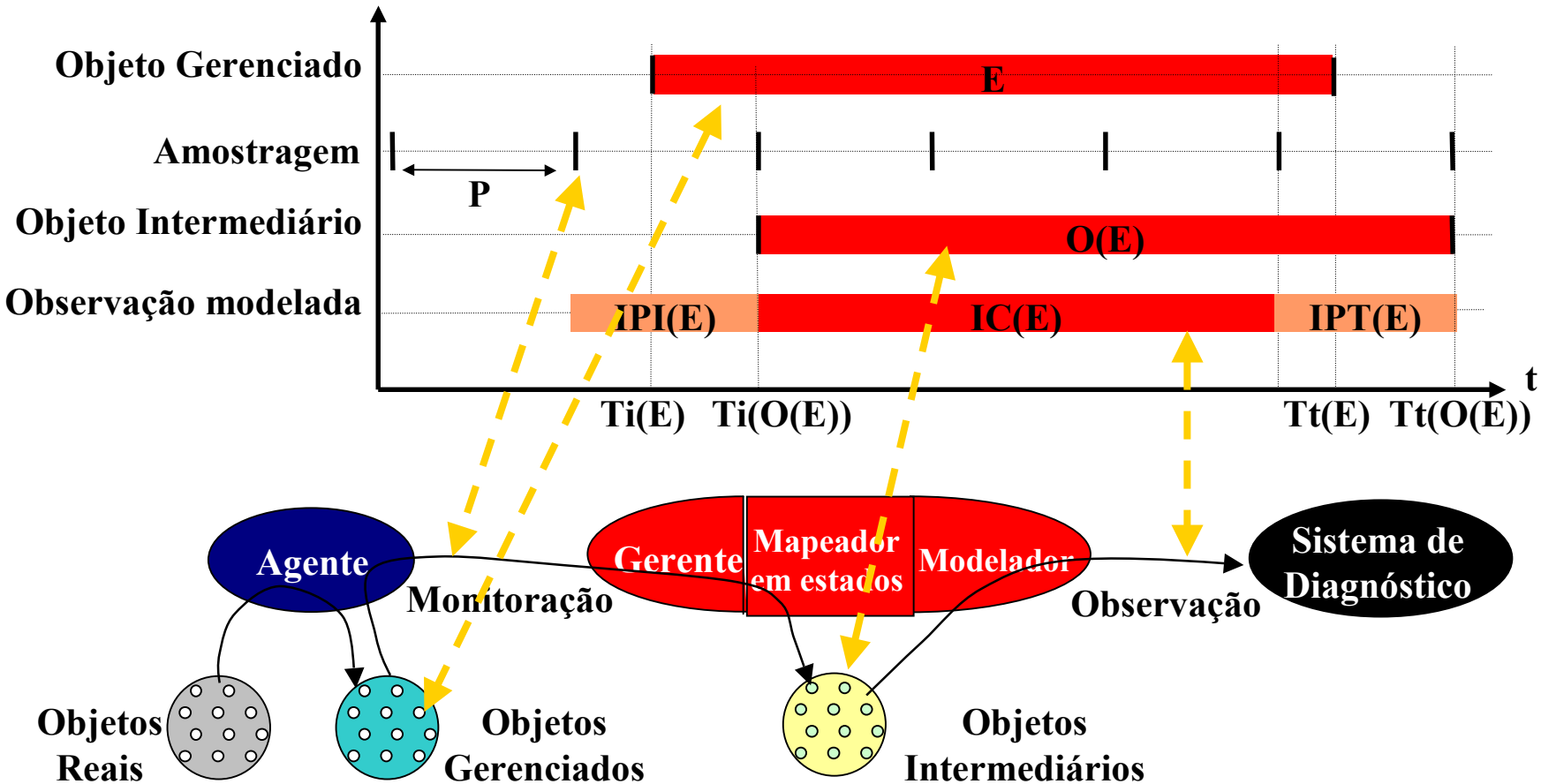
Intervalos de possibilidade e certeza

- Observação não defasada



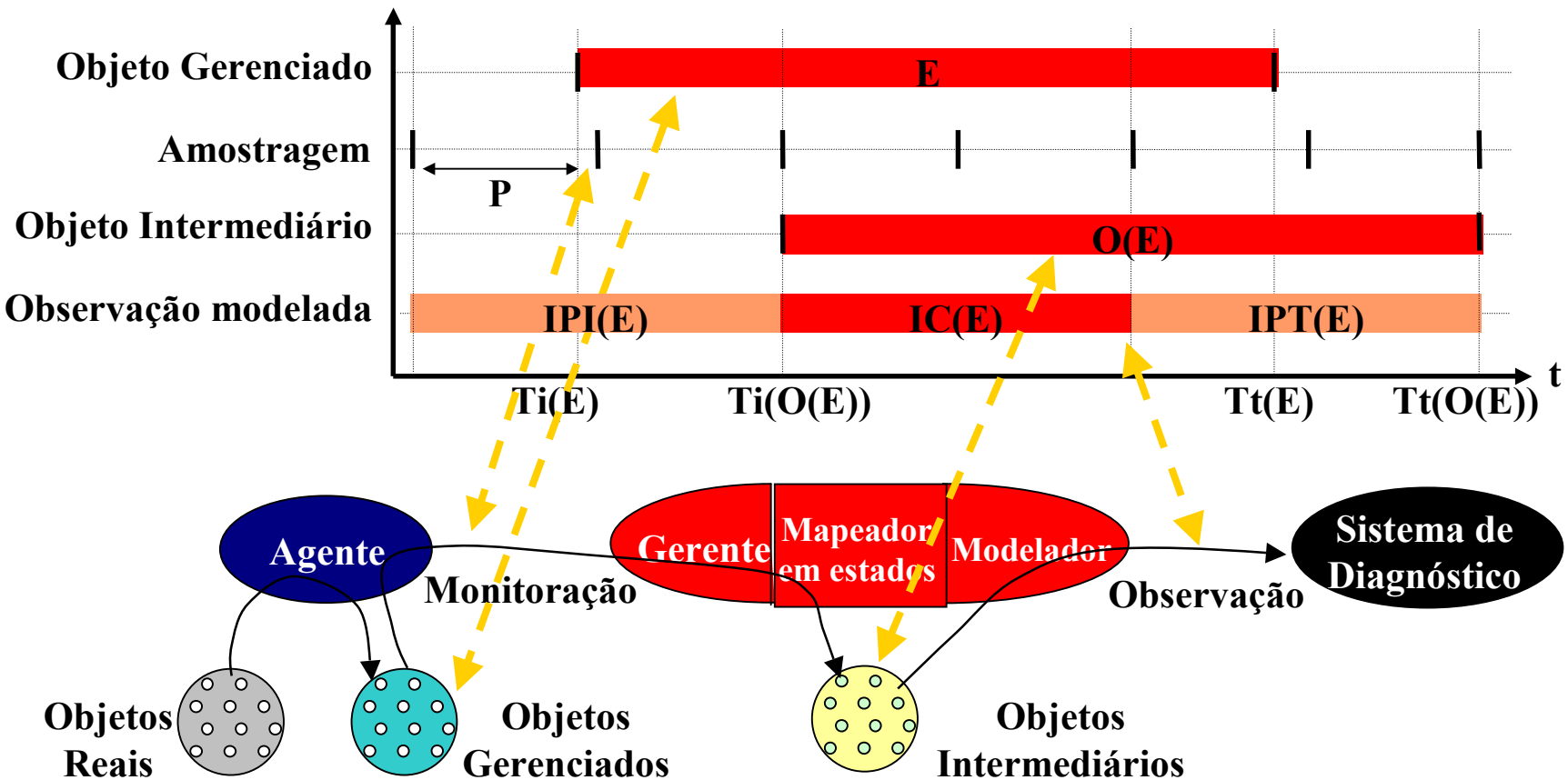
Intervalos de possibilidade e certeza

- Observação defasada em até 1 ciclo



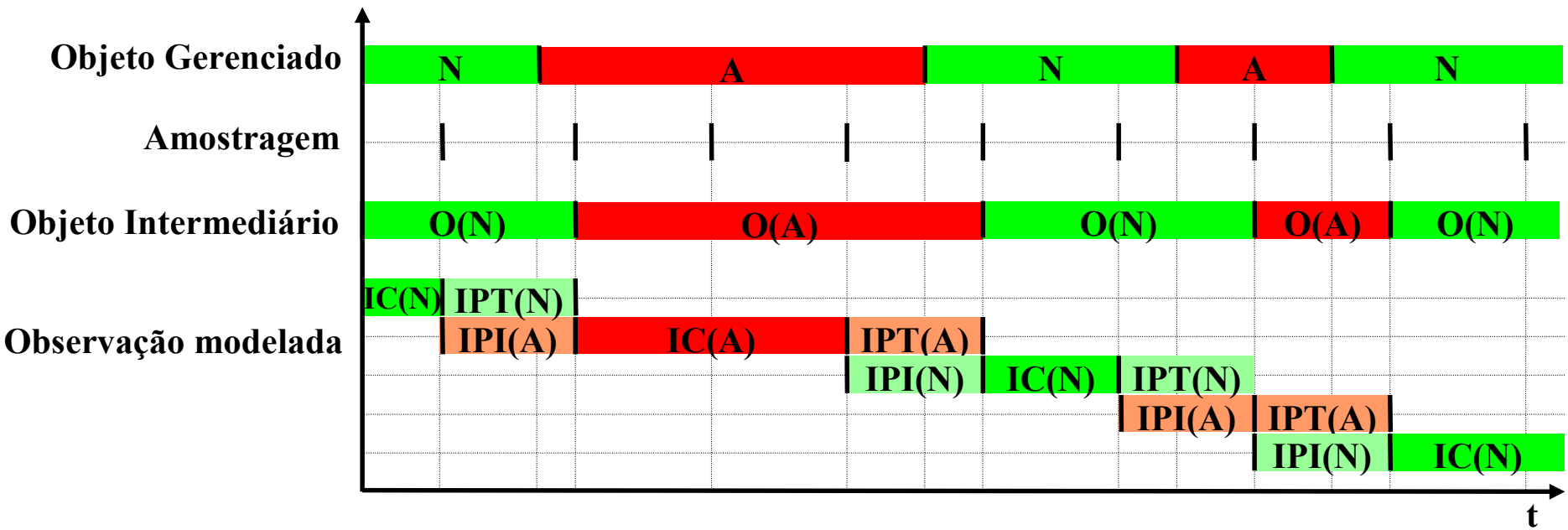
Intervalos de possibilidade e certeza

- Observação defasada em até 2 ciclos



Intervalos de possibilidade e certeza

- Utilização da completude do conjunto de estados



Modelagem da observação

Intervalos de Incerteza

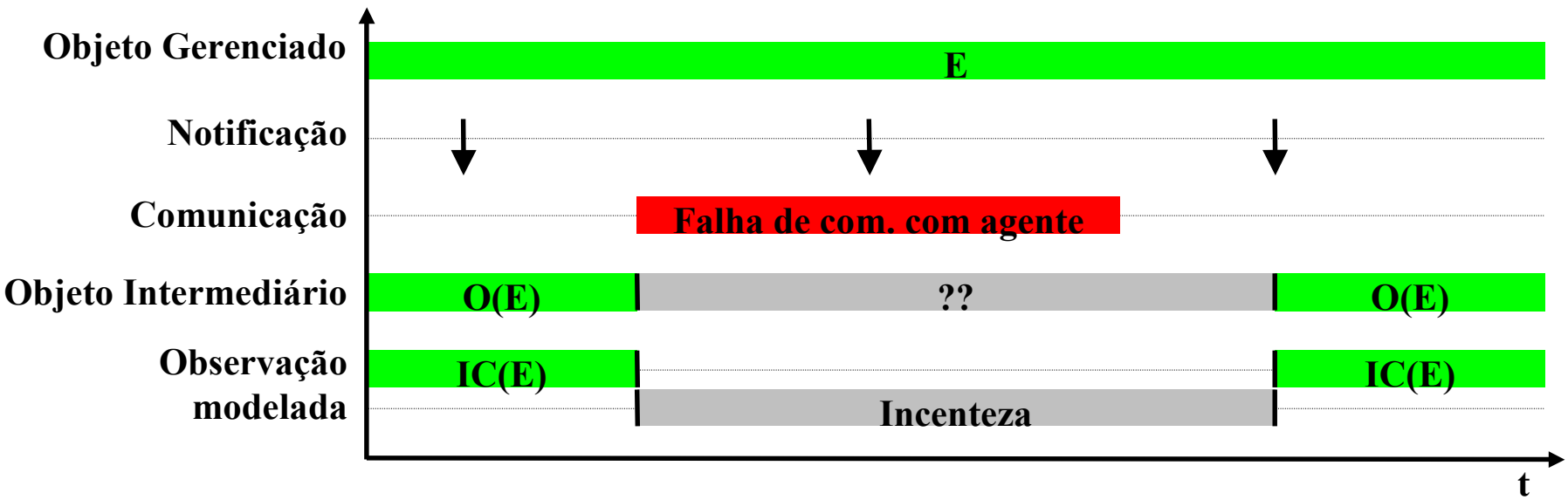
Intervalos de Incerteza

- Representa o intervalo de tempo no qual não é conhecido o "estado" do objeto gerenciado

- Causas
  - (1) Falha na monitoração
  - (2) Proximidade ao instante corrente
  - (3) Início do processo de monitoração

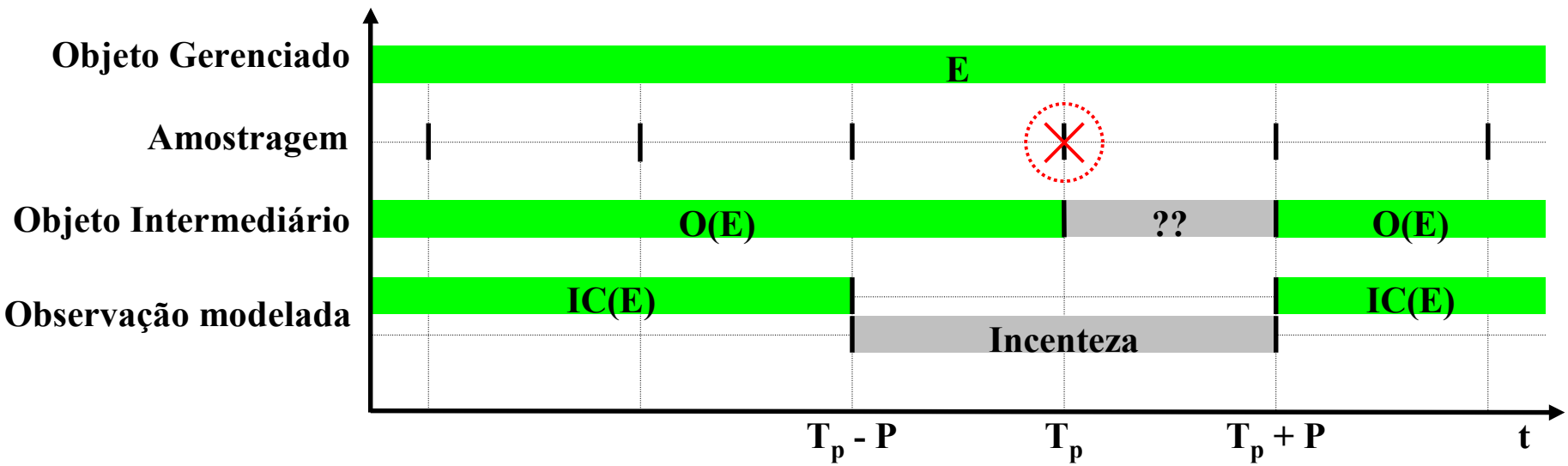
Intervalos de incerteza: (1) Falha na monitoração

- Observação não defasada



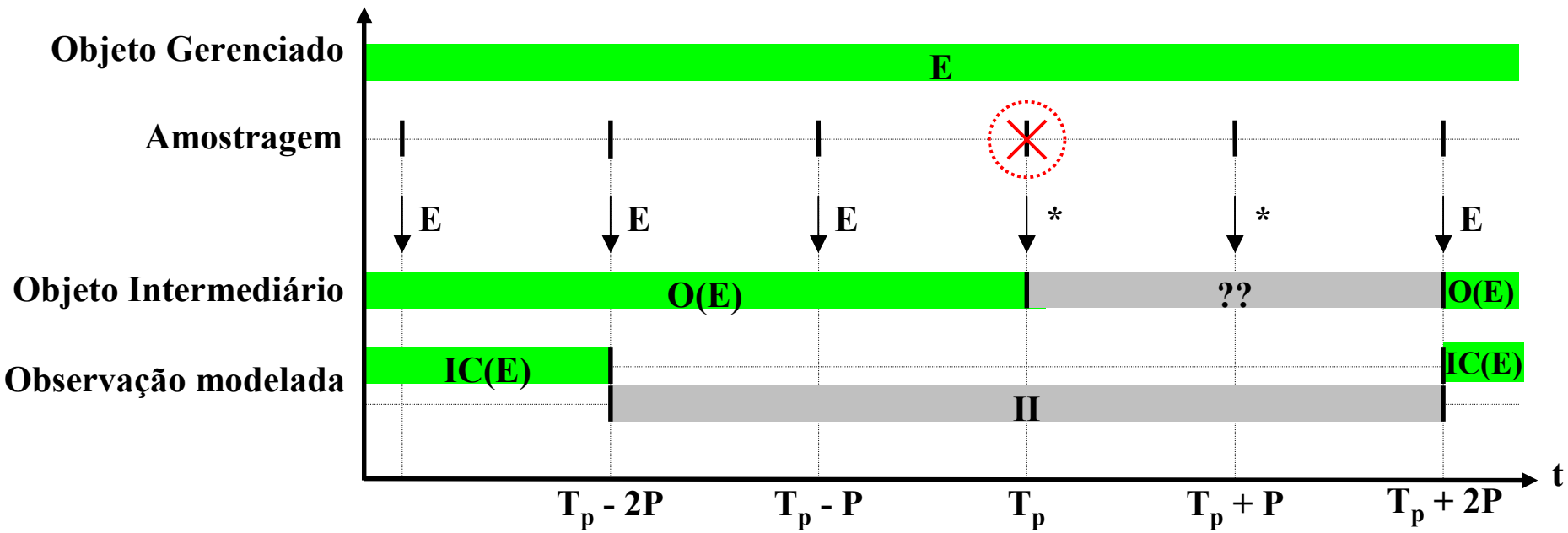
Intervalos de incerteza: (1) Falha na monitoração

- Observação defasada em até 1 ciclo



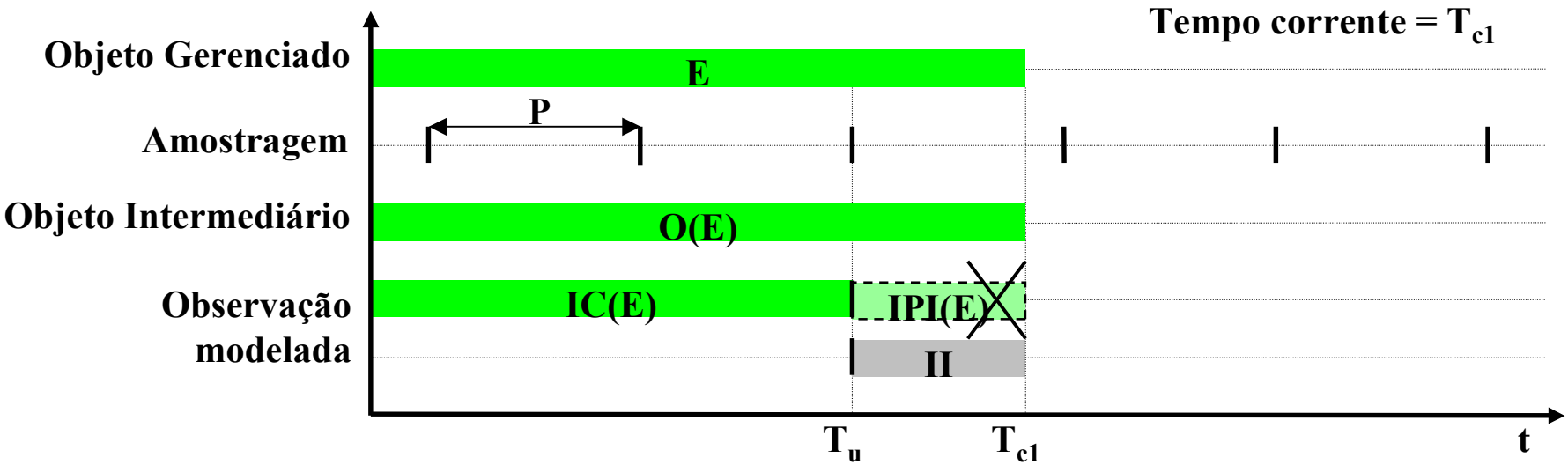
Intervalos de incerteza: (1) Falha na monitoração

- Observação defasada em até 2 ciclos



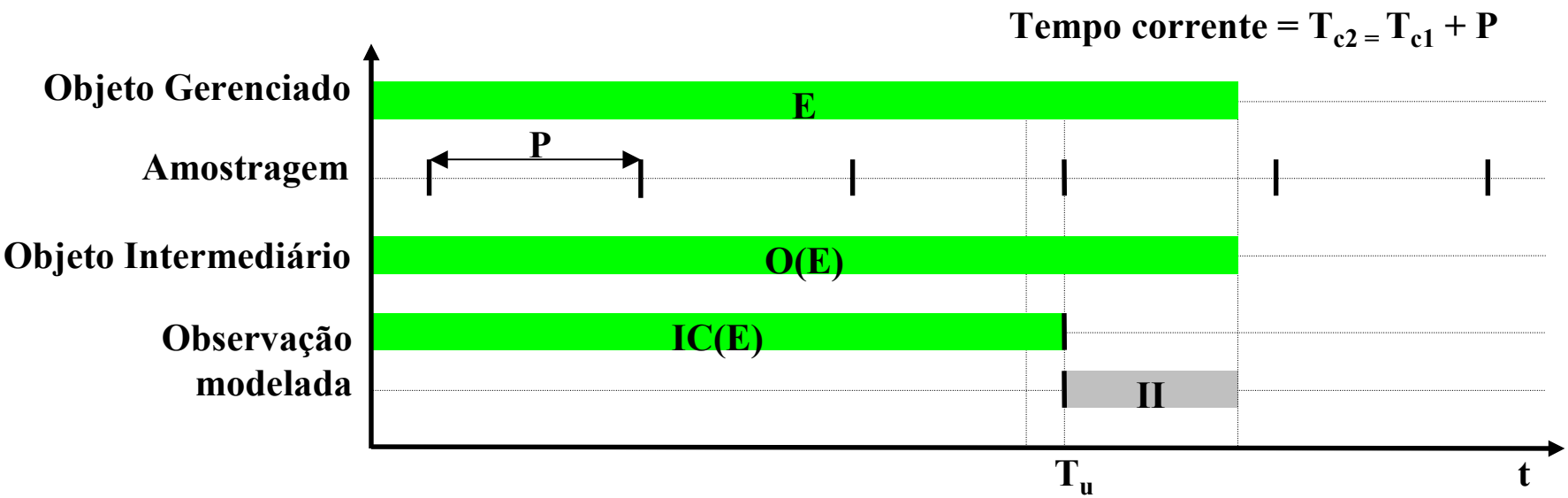
Intervalos de incerteza: (2) Proximidade ao instante corrente

- Observação defasada em até 1 ciclo



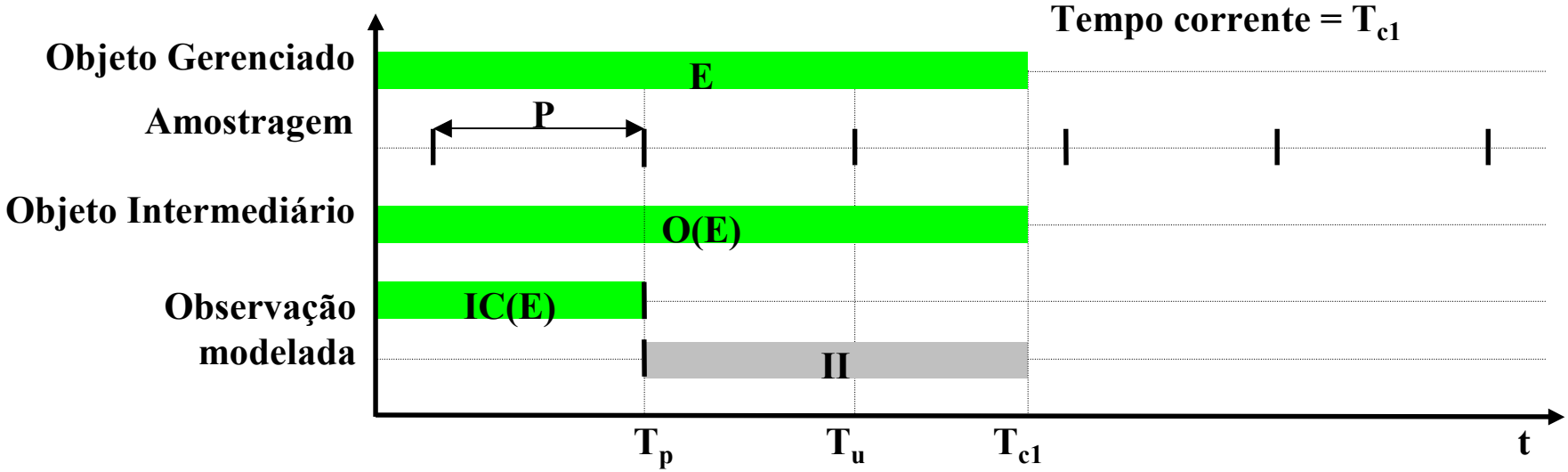
Intervalos de incerteza: (2) Proximidade ao instante corrente

- Observação defasada em até 1 ciclo



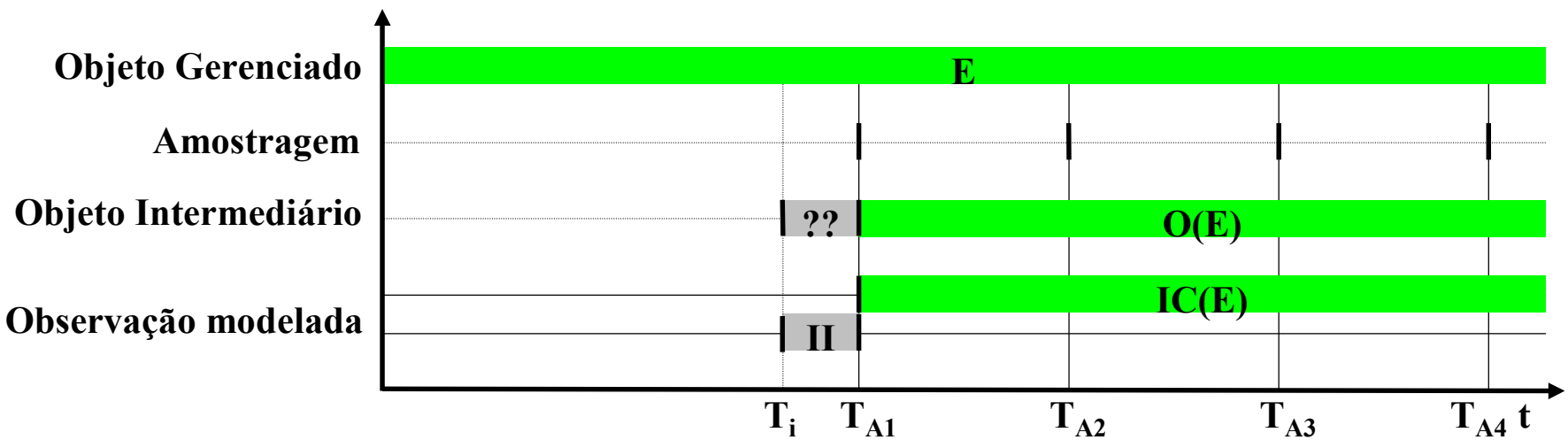
Intervalos de incerteza: (2) Proximidade do instante corrente

- Observação defasada em até 2 ciclos



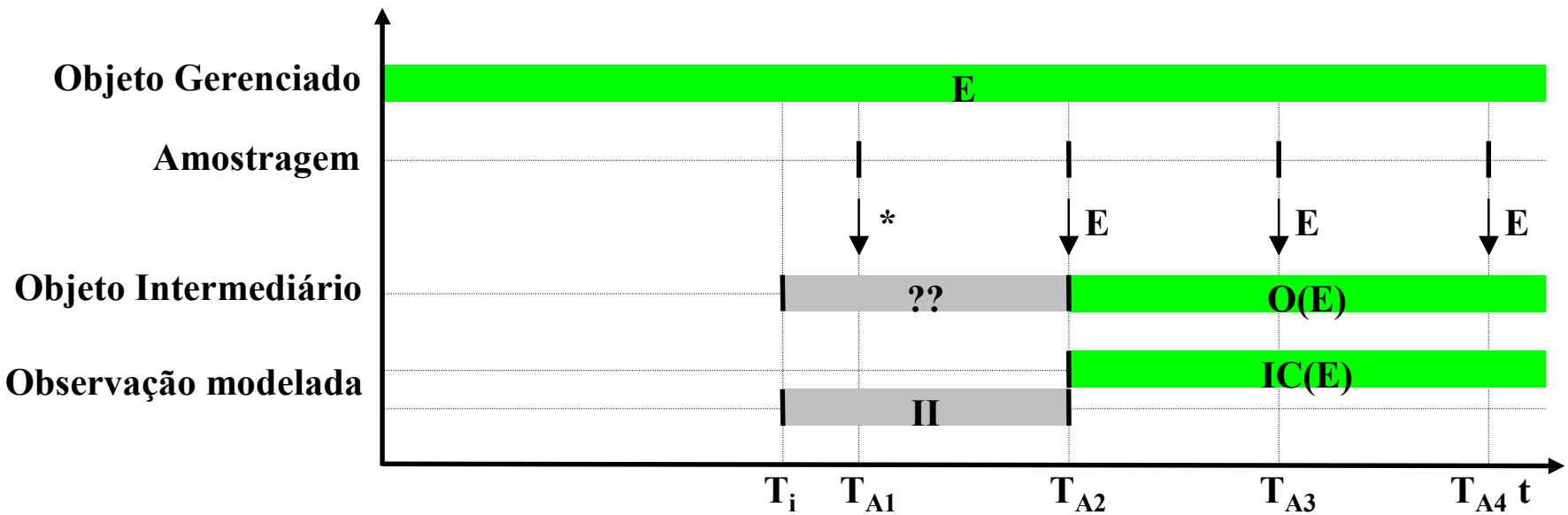
Intervalos de incerteza: (3) Início do processo de monitoração

- Observação defasada em até 1 ciclo



Intervalos de incerteza: (3) Início do processo de monitoração

- Observação defasada em até 2 ciclos



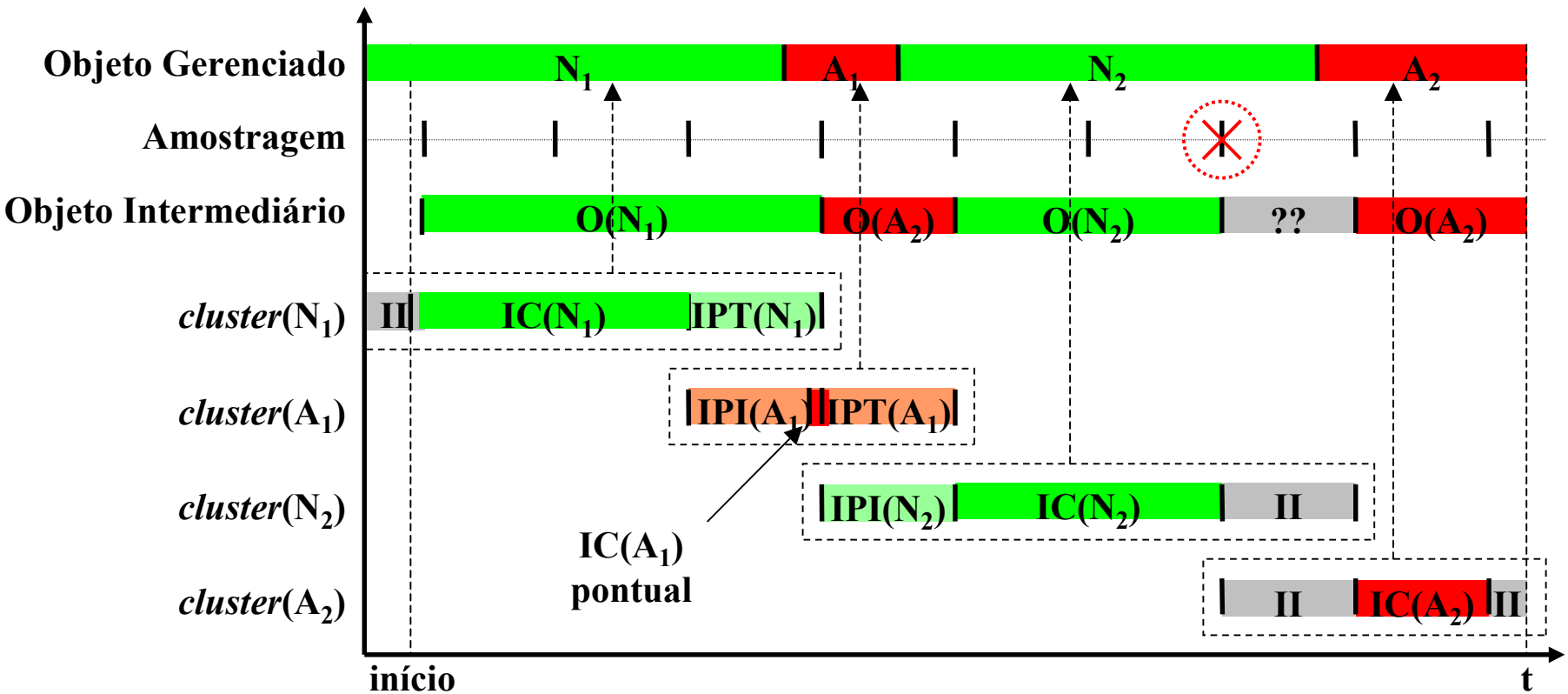
Aglomerado de Intervalos

Aglomerado de intervalos

- Objetivo
  - Representar a ocorrência de um estado em um objeto gerenciado utilizando os intervalos de tempo IPI, IC, IPT e II

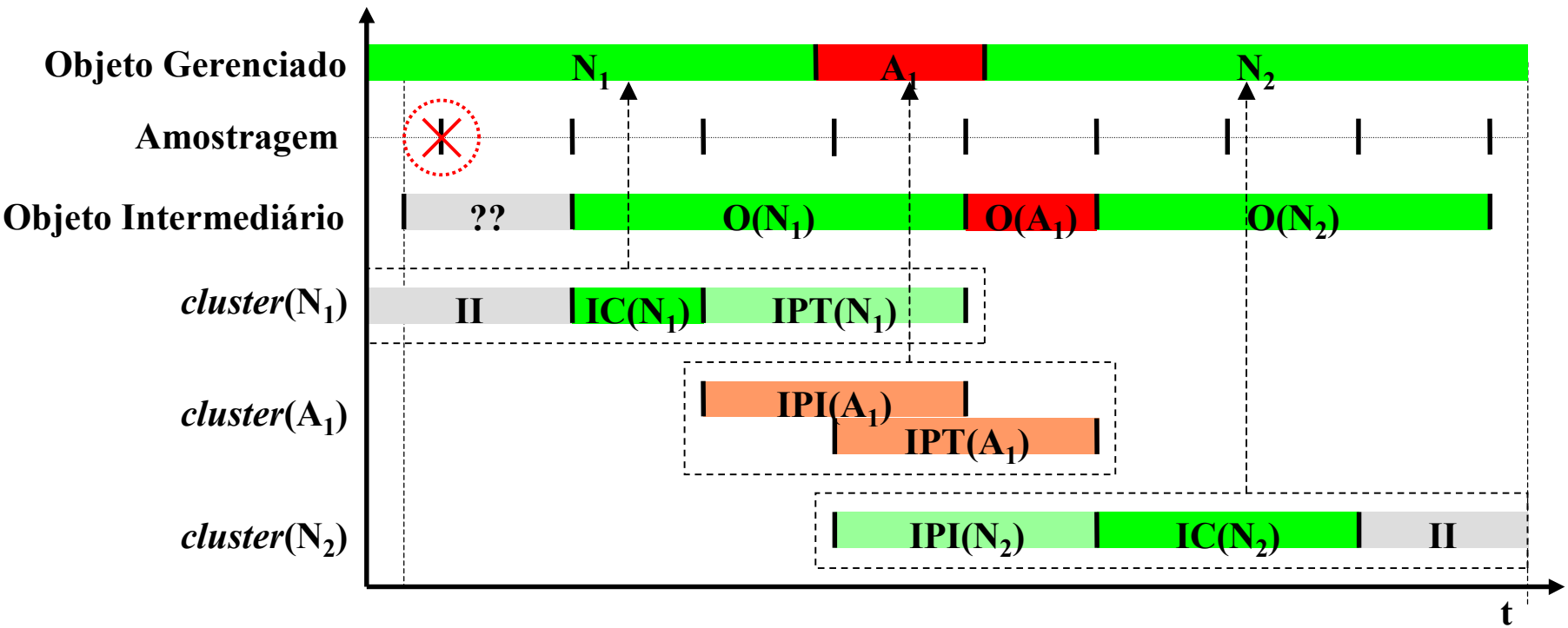
Aglomerado de intervalos

- Exemplo:
  - Observação defasada em até 1 ciclo



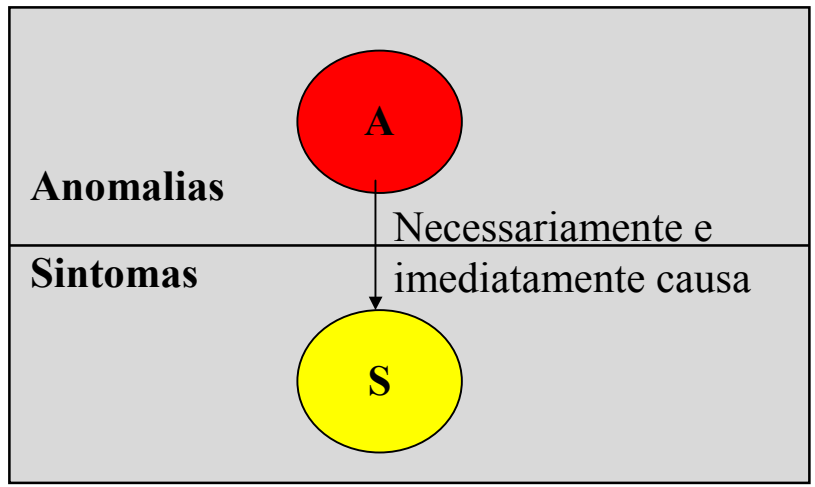
Aglomerado de intervalos

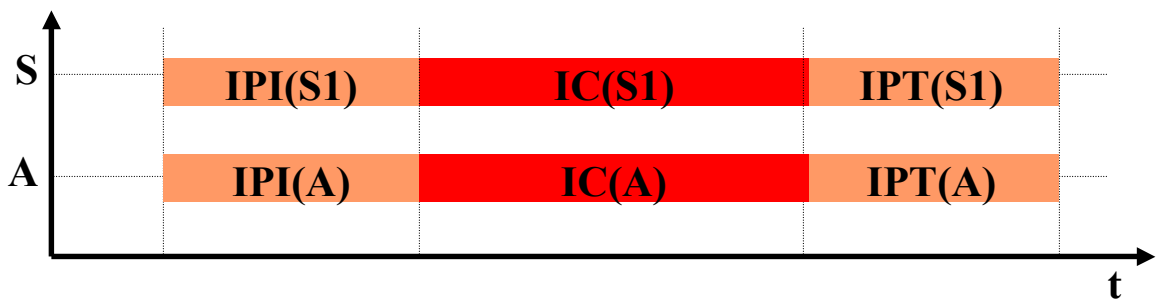
- Exemplo:
  - Observação defasada em até 2 ciclos



Aglomerado de intervalos

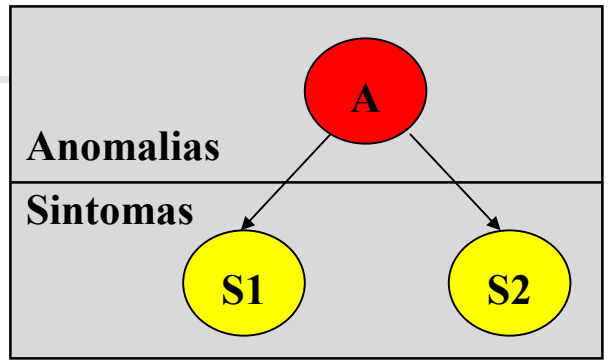
- Inferência

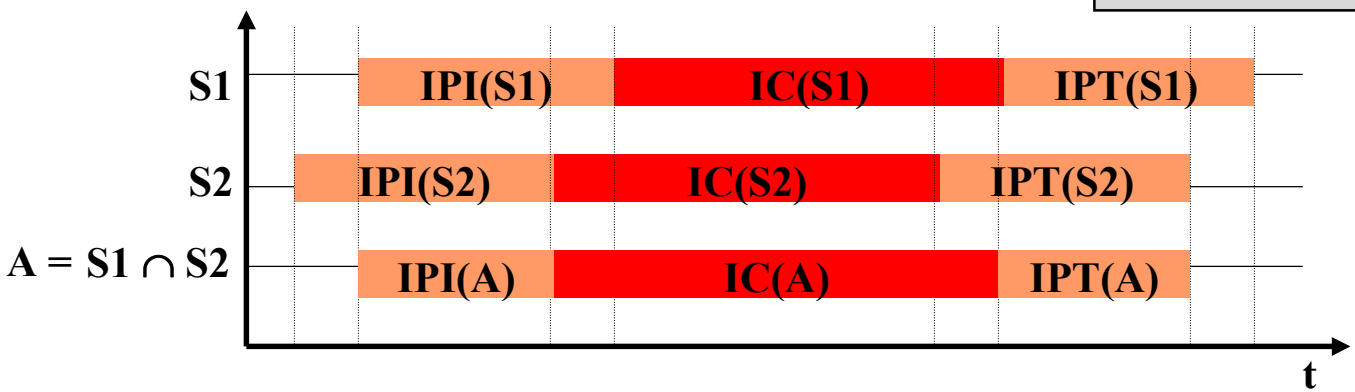


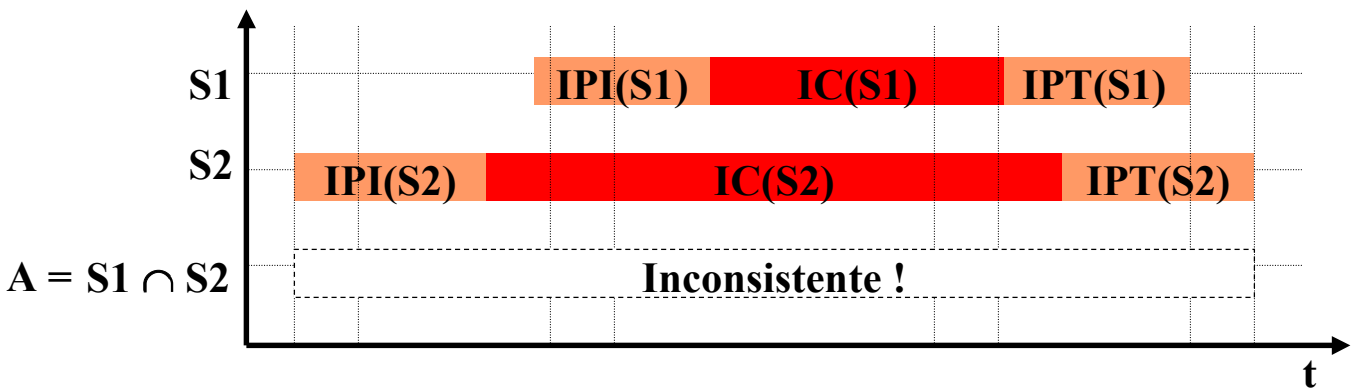


Aglomerado de intervalos

- Intersecção entre aglomerados







Conclusão

Conclusão

- Redes de dados possuem características distintas das redes de telecomunicações em relação ao gerenciamento
- Técnicas de correlação, empregadas extensivamente em redes de telecomunicações não são tão eficientes em redes de dados
- Informações temporais são descartadas em sistemas tradicionais de diagnóstico

Conclusão

- Informação temporal contida na observação pode ser utilizada para discriminar hipóteses em um sistema de diagnóstico. Isto é útil
  - Quando observações possuem diferentes períodos e fase de amostragem
  - Quando da presença de múltiplas anomalias

- Porém, em redes de dados, a informação temporal contida em uma observação é imprecisa. Daí a importância de explicitar tais imprecisões na observação.

Conclusão

- Este trabalho
  - Modelou o processo de obtenção de informações em um sistema distribuído
  - Definiu novos termos e propriedades relacionados à observação, não encontrados na literatura
  - Propôs uma forma de modelagem da observação, explicitando suas imprecisões temporais
  - Exemplificou como estas observações podem ser utilizadas em um sistema de diagnóstico temporal

Conclusão

- Observação gerada por sistemas tradicionais possui o problema da imprecisão temporal, intrínseco à dinâmica do processo de observação
- Proposta de modelagem da observação:
  - Vantagens
    - Pode ser utilizado em sistemas de diagnóstico temporal e atemporal
    - Permite ao sistema de diagnóstico o conhecimento das imprecisões temporais e incertezas de observação
  - Desvantagem
    - Processo deve estar integrado ao sistema coletor
- A modelagem da observação é o primeiro passo para concepção de um sistema de diagnóstico temporal

Contribuições

Contribuições

- Modelagem do processo de observação em sistemas distribuído
  - Definição de objeto intermediário
  - Tipos de observação
  - Classificação das observações em
    - Não defasadas
    - Defasadas em até 1 ciclo
    - Defasadas em até 2 ciclos

Contribuições

- Proposta de modelagem temporal da observação
  - Definição de uma forma de observação que incorpora informação a respeito da imprecisão temporal (IPI, IC, IPT, II)
- Modelos reusáveis para a representação de um sistema distribuído [Bernal 1999b]
- No grafo causal, foi explicitando:
  - O plano de observações
  - O plano de anomalias

Trabalhos Futuros

Trabalhos Futuros

- Modelagem da observação: formalização matemática de agrupamento de intervalos e operações sobre tais agrupamentos
- Pesquisa de outros métodos de diagnóstico temporal
- Comparação efetiva de sistemas de diagnóstico atemporal e temporal em relação a velocidade e precisão de diagnóstico
- Uso desta técnica de modelagem em um sistema de produção: diagnóstico de falhas em um cluster de processamento paralelo (em andamento)

Obrigado