From Data-Centric Business Processes to Enterprise Process Frameworks

Jianwen Su
UC Santa Barbara

A Traveler's Experience

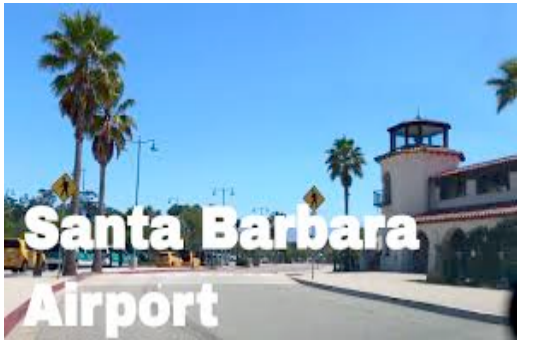
friendly sky…

got a phone…

tax refund…

Immigration & Border Protection

back home



return the phone

upgrade

Immigration & Border Protection

tax refund of a new phone

Business Services via an Example

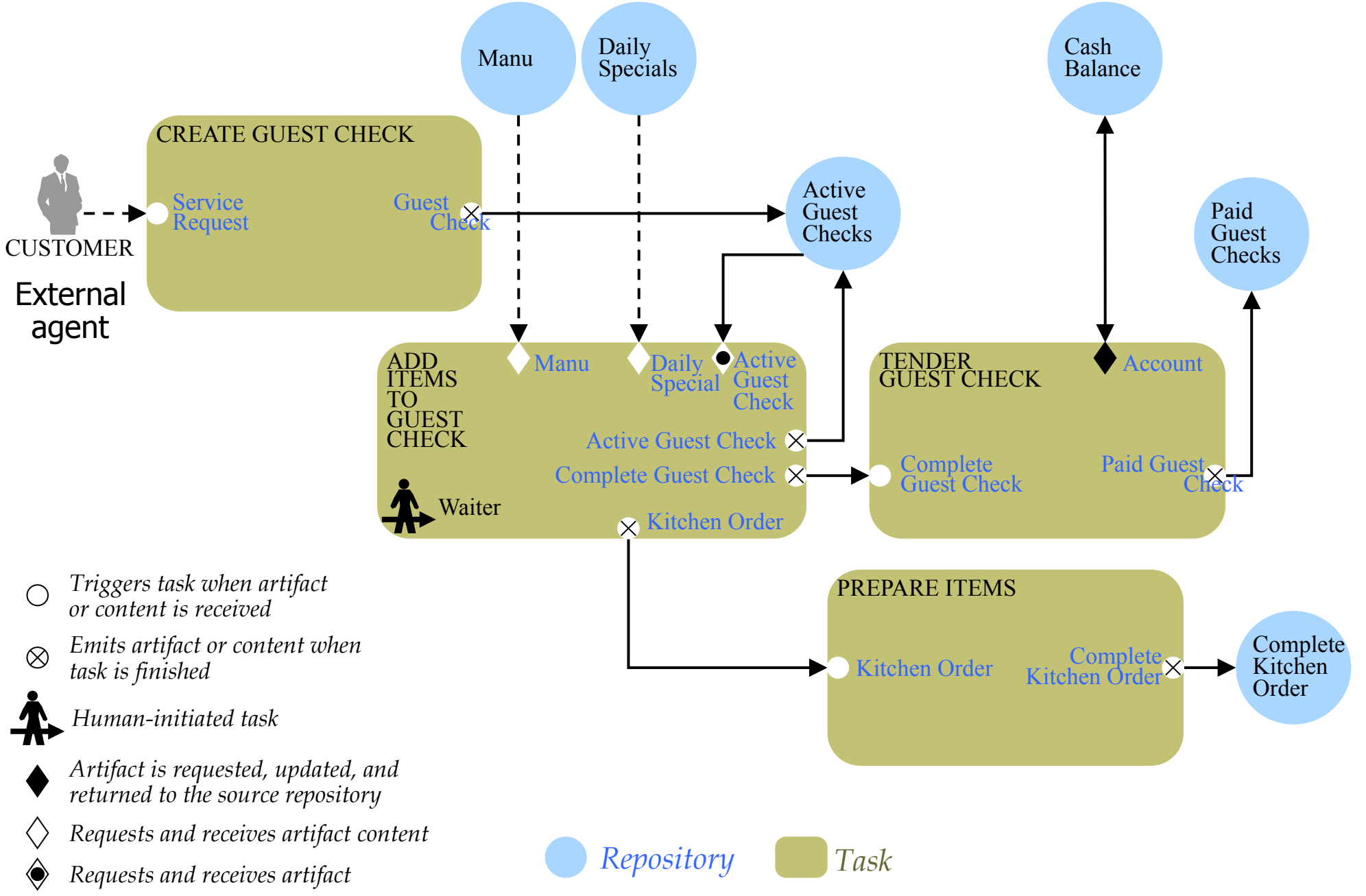
- Subscription of broadband internet connection
- A collection of processes: *order*, *shipping*, *installation*, *invoice*, *cancel*, *update order*, …
  - Usually well specified and (formally) modeled
- Relationships between processes: *order* triggers *shipping* and *installation*, *installation* triggers *invoice*, *cancel* triggers *invoice*, …
  - Usually in biz rules, documents, and manual
- Properties of relationships:
  - Often data-centered
  - Broader than choreography
- Modeling processes + relationships is beneficial

Plan for the Talk

- The Need for Process-Process Relationships
- Process Design and Modeling with Data
- Runtime Management
- Towards Process-Relationship (PR) Modeling
- Further Challenges

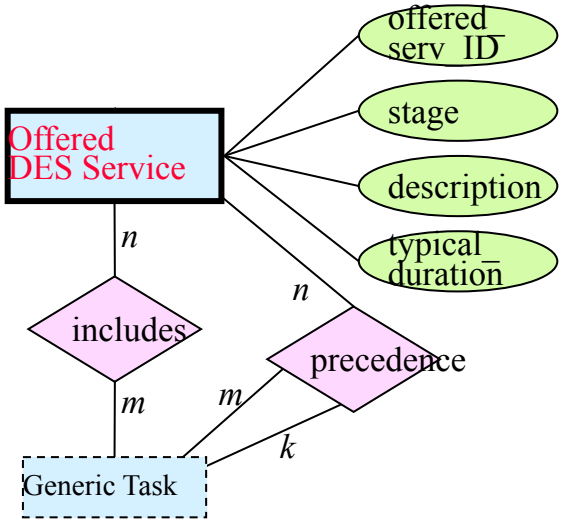
Life Cycle of Guest Check Artifact

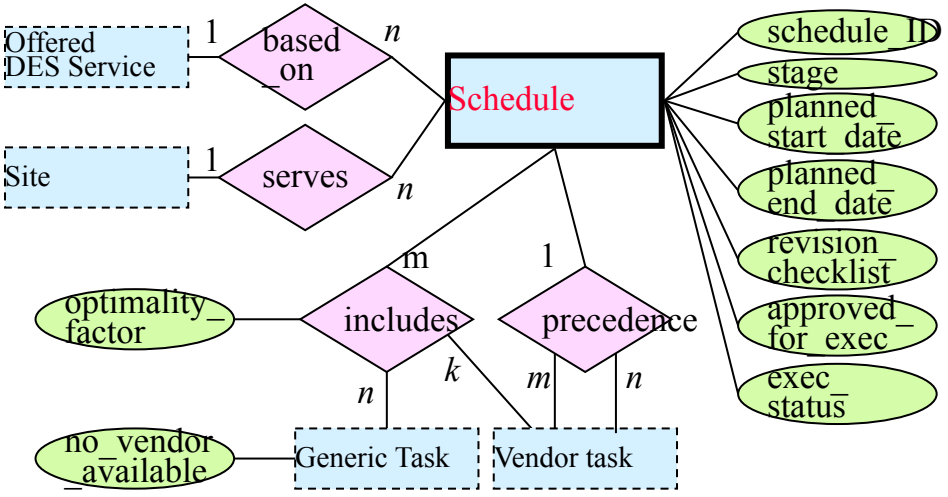
[Nigam-Caswell 03]

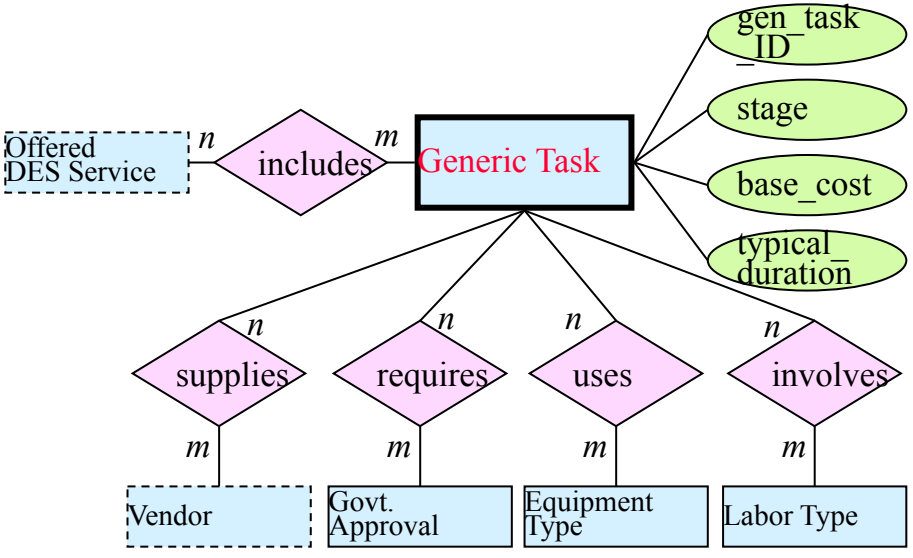


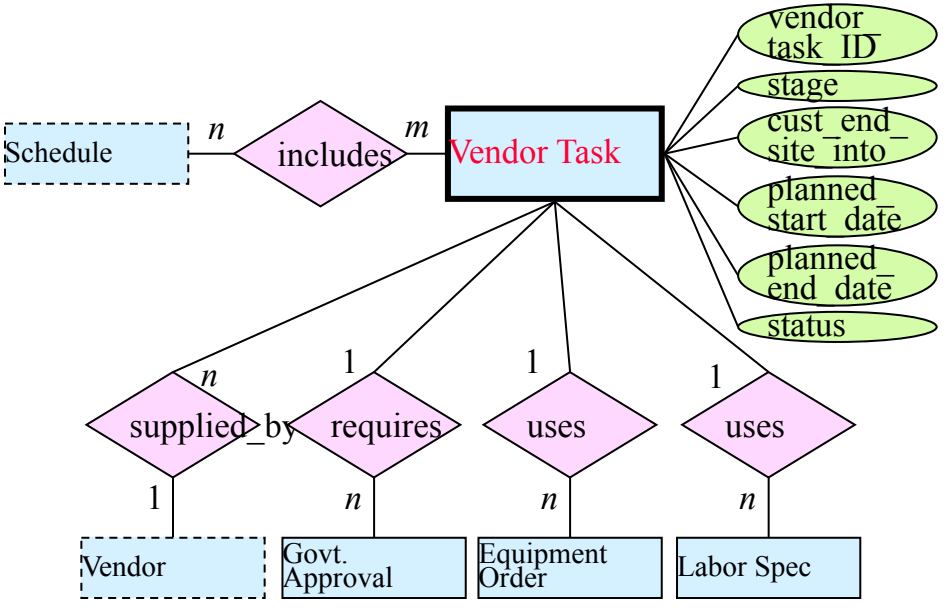
Discovery and Design of Artifacts

- ER diagrams or other suitable modeling approaches





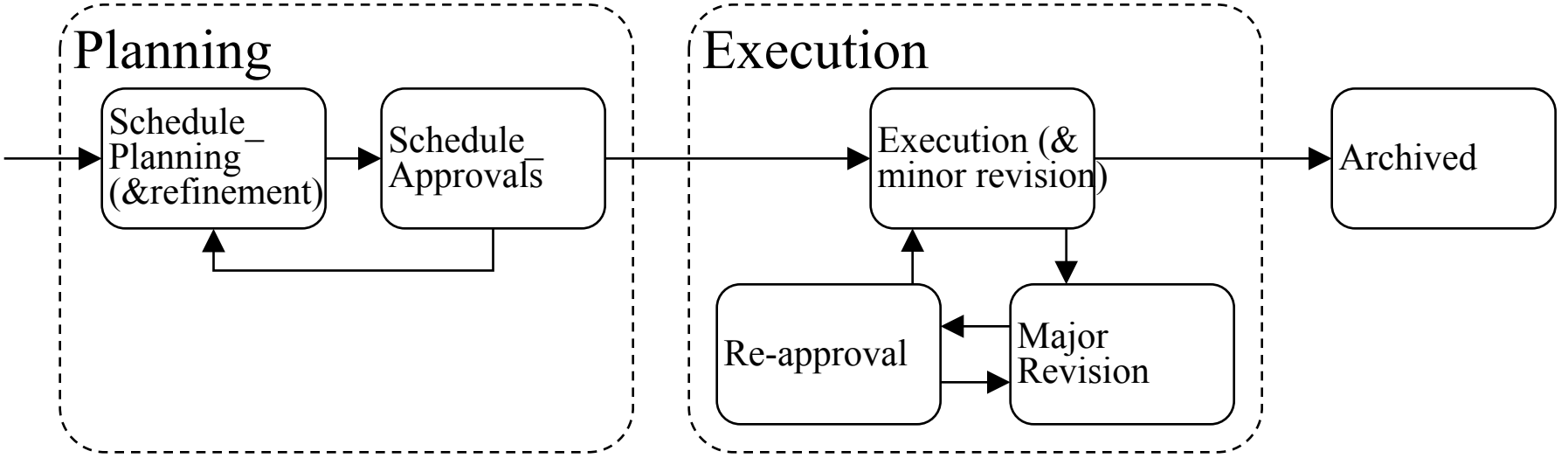




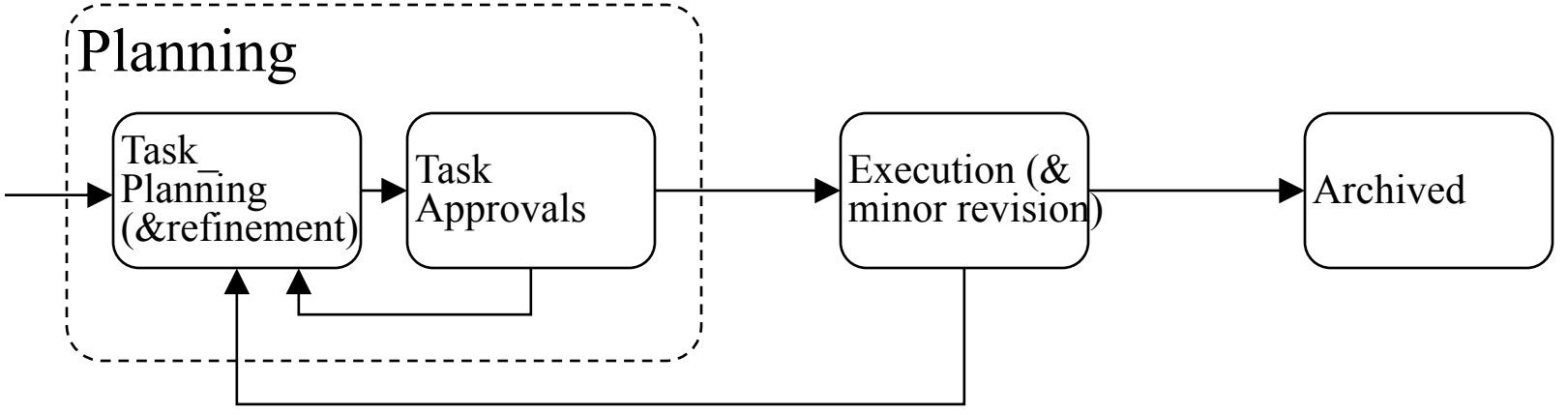
 [Bhattacharya-Gerede-Hull-Liu-S. BPM 07] [Bhattacharya-Hull-S. 09]

Schedule and Vendor Lifecycles

- Schedule



- Vendor



BOM Service: IOPEs of *Create_schedule*

Inputs

- An *Offered DES Service* artifact *o*, and specifically the listing of used *Generic Tasks*, along with whether they are optional, and information about the Precedence relationships between them
- A *Customer* artifact *c*, ...
- A *Site* artifact *si* for *c*, ...

Outputs

- A new *Schedule* artifact *sch*. The data written will include attributes *schedule_ID*, *stage*, *planned_start_date*, and the *Generic Task* portion of the *includes* relationship
- The *Site* artifact *si* is updated ...

Pre-

- *Offered DES Service* artifact *o* must be compatible with the infrastructure and needs of site *si*

Cond. effect

- If true, then *sch* is in stage *Schedule_planning*
- If true, then *sch* holds a schedule skeleton (i.e., appropriate portions of the relationship *includes* are filled in)
- If true, ...

BOM: ECA Rules

R_1: *initiate schedule*

event request by performer p to create a schedule instance for *Offered DES Service* artifact o, *Customer* artifact c, and *Site* artifact si

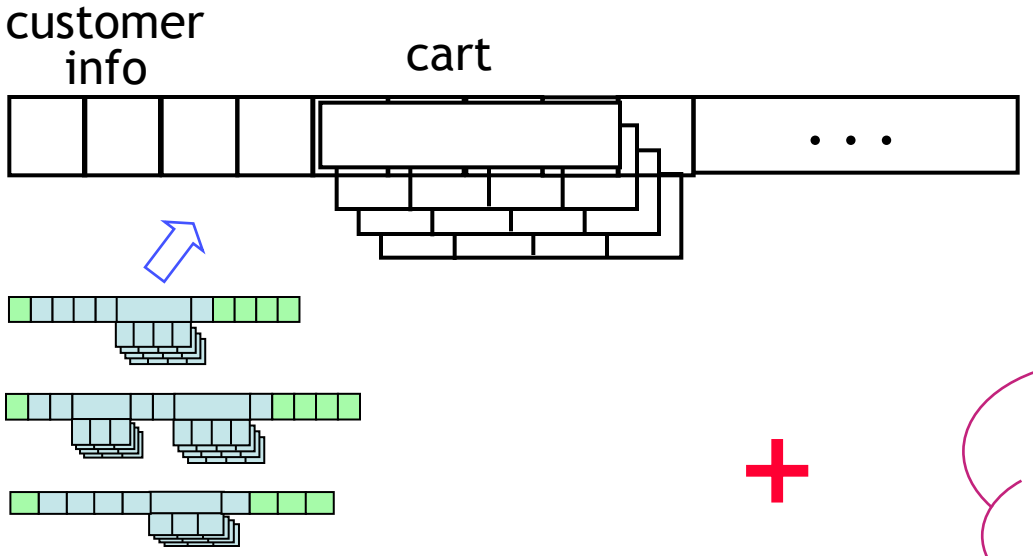
condition the appropriate non-disclosure agreements (NDAs) are in place for c

action invoke *Create_schedule*(o, c, si)

by performer p where *offer_manager in role*(p) and *qualification*(p, o, region: $si.region$) ≥ 5

Alternative models can also be used

Artifact-Centric Biz Process Modeling



Artifacts (Info models)

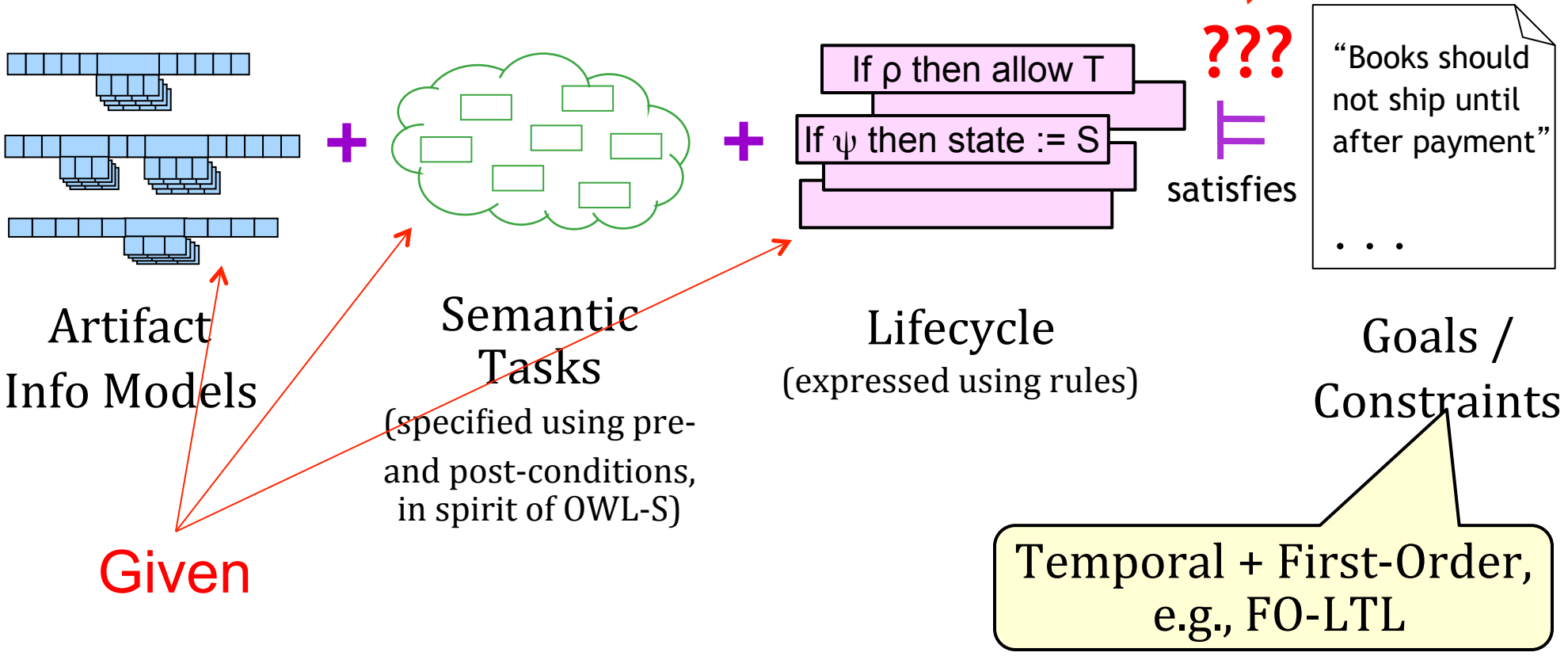
+

Specification of artifact lifecycles

- Informal model [Nigam-Caswell 03]
- Systems: BELA (IBM 2005), Siena (IBM 2007), ArtiFlow, EZ-Flow (Fudan-UCSB 2010), GSM/Barcelona (IBM 2010)
- Formal models
  - State machines [Gerede-Bhattacharya-S. SOCA07][Gerede-S. ICSOC07]
  - Rules [Bhattacharya-Gerede-Hull-Liu-S. BPM07] [Hull et al WS-FM10]

Design Analytics

- Workflow verification problem:



- An important problem [Hull-S. DCW09 report]
- More in SIGMOD tutorial [Hull-S.-Vaculin SIGMOD13]

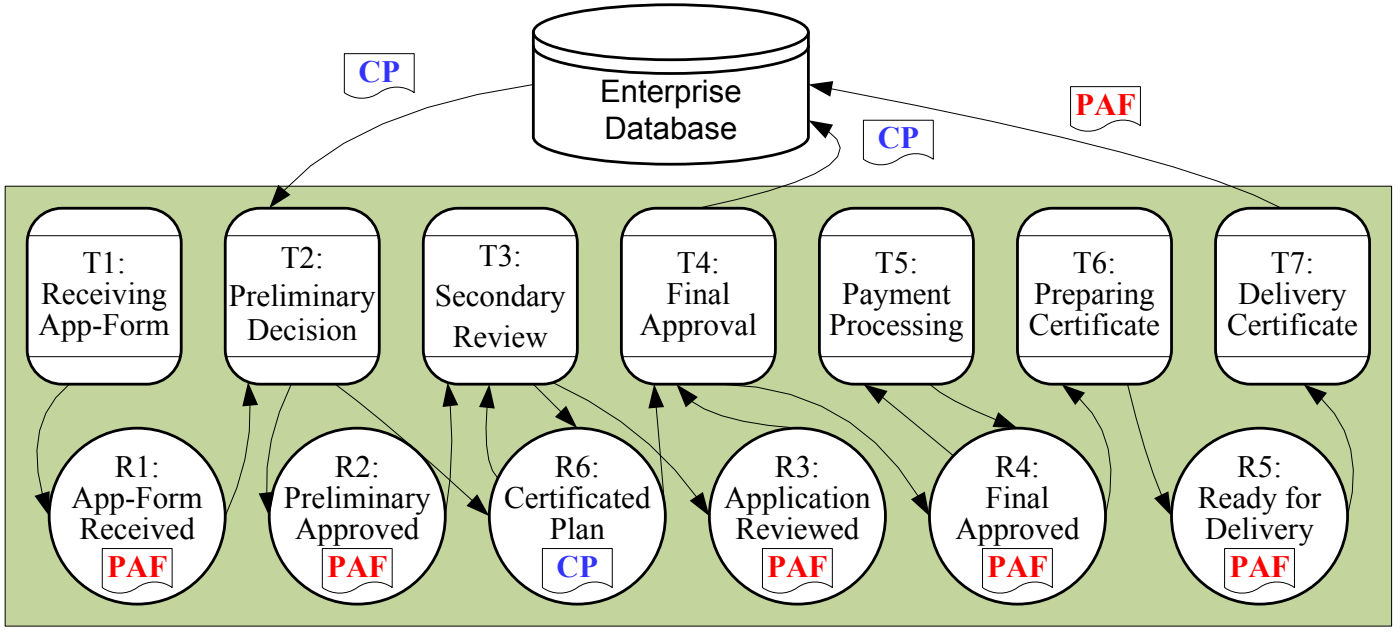
Plan for the Talk

- The Need for Process-Process Relationships
- Process Design and Modeling with Data
- Runtime Management
- Towards Process-Relationship (PR) Modeling
- Further Challenges

Dynamic Process Changes

[Xu.-S.-Yan-Yang-Zhang CoopIS11]

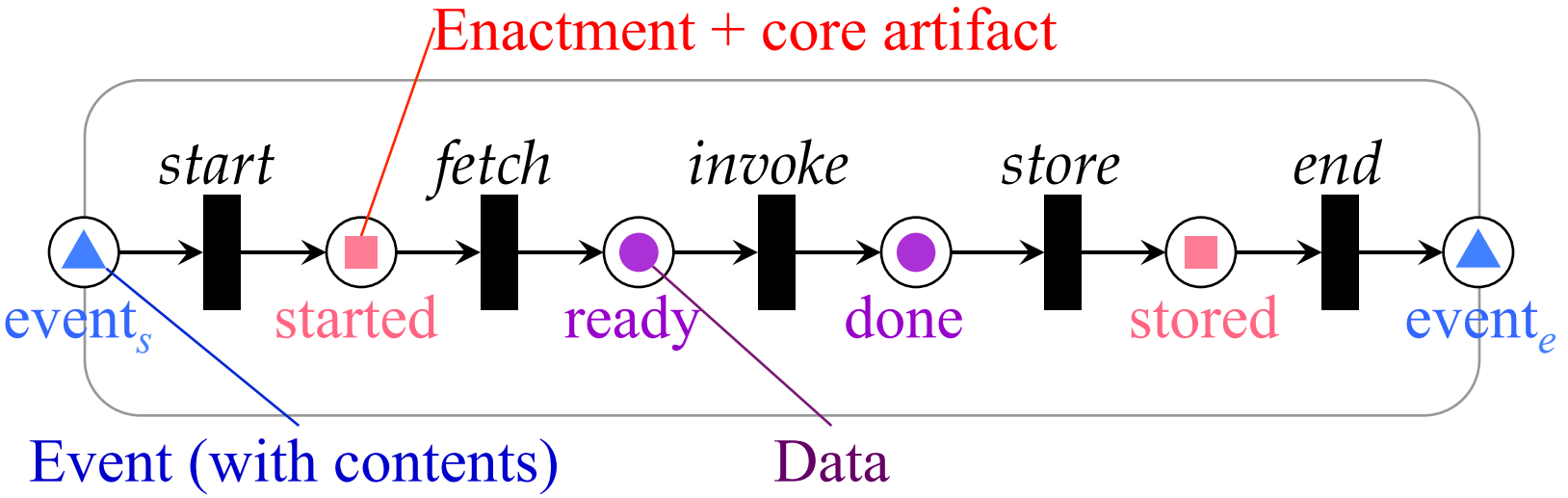
- Artifact-centricity, EZ-Flow model



- Each biz process has a core artifact (class)
  - Business data (object) + enactment
  - Similar notion in recent GSM model from IBM

Execution Semantics and Process Changes

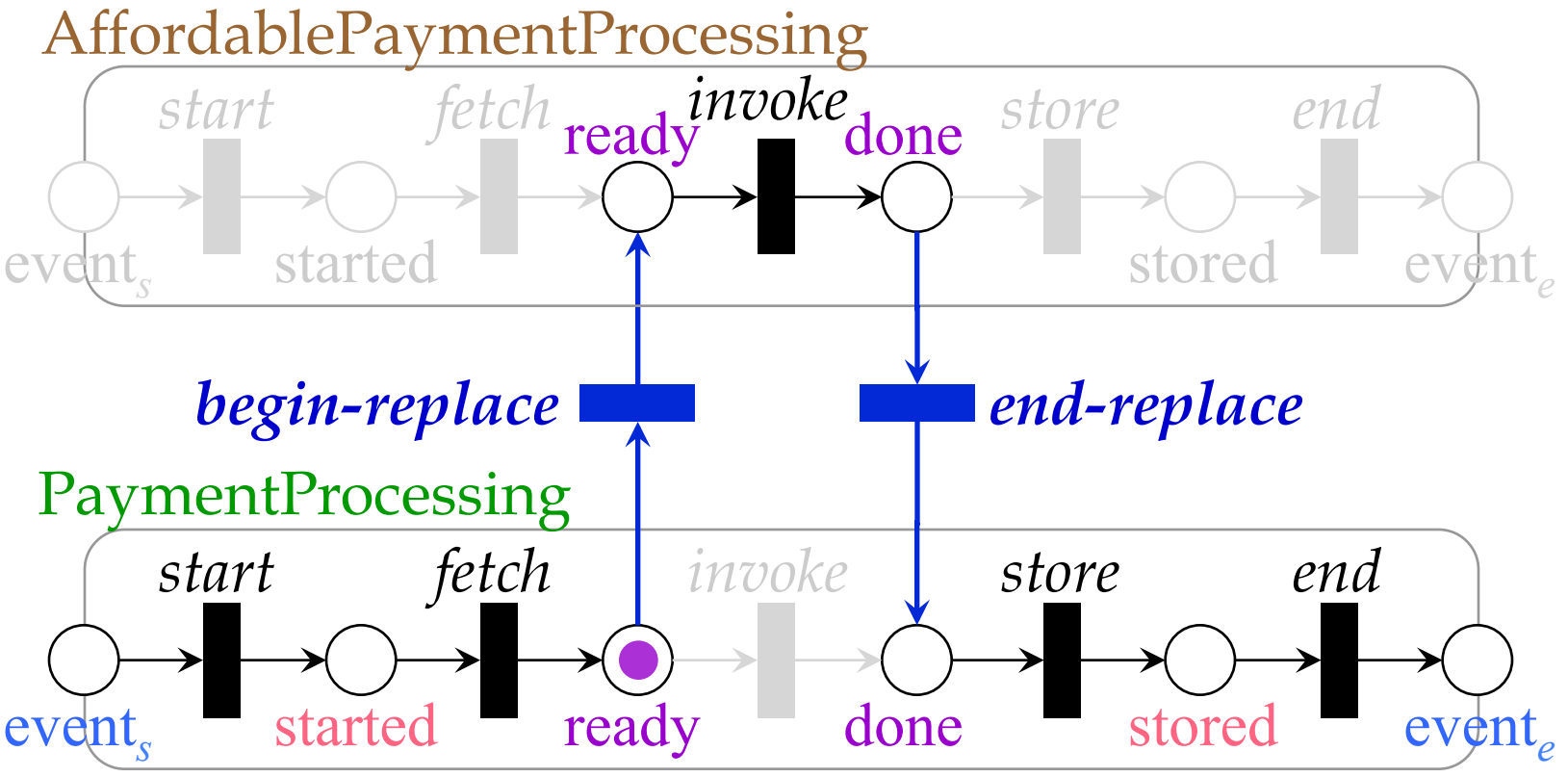
- Formal semantics for task execution based on Petri nets



- Represents data (input/output) requirements and carries enactments
- Declarative change specification
  - Four execution altering operators
  - Rules for applying the operators based on conditions

[Xu.-S.-Yan-Yang-Zhang CoopIS11]

New Fee Schedule for Low Income Housing




```
Affordable-Fee:
  MUST REPLACE PaymentProcessing
               BY AffordablePaymentProcessing ON PAF
WHERE SELF.projectType="affordable"
```

[Xu.-S.-Yan-Yang-Zhang CoopIS11]

Jointfounder Challenge

杭州中房信息科技有限公司
Hangzhou Jointfounder InformationTechnology Co., Ltd

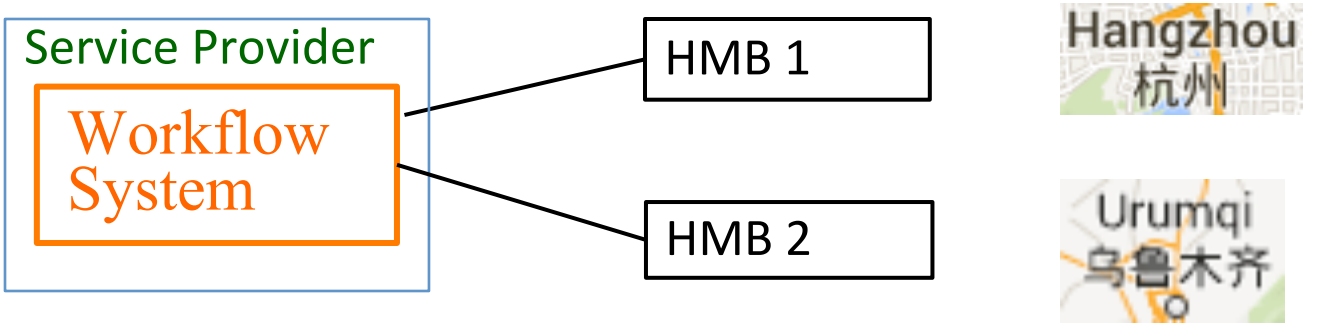
- Housing Management Bureau (HMB房管局) manages *titles*, *licenses*, *permits*, ... for a region
- Sells housing management workflow systems to HMBs
  - 20-30 HMBs as clients, including 

- Maintenance contracts for clients
  - Each service call costs 4 - 6 person-days
  - Common types of issues:
    failures, changes caused by e.g. policy change,
    (tools for) analytics, ...
- Scalability problem:

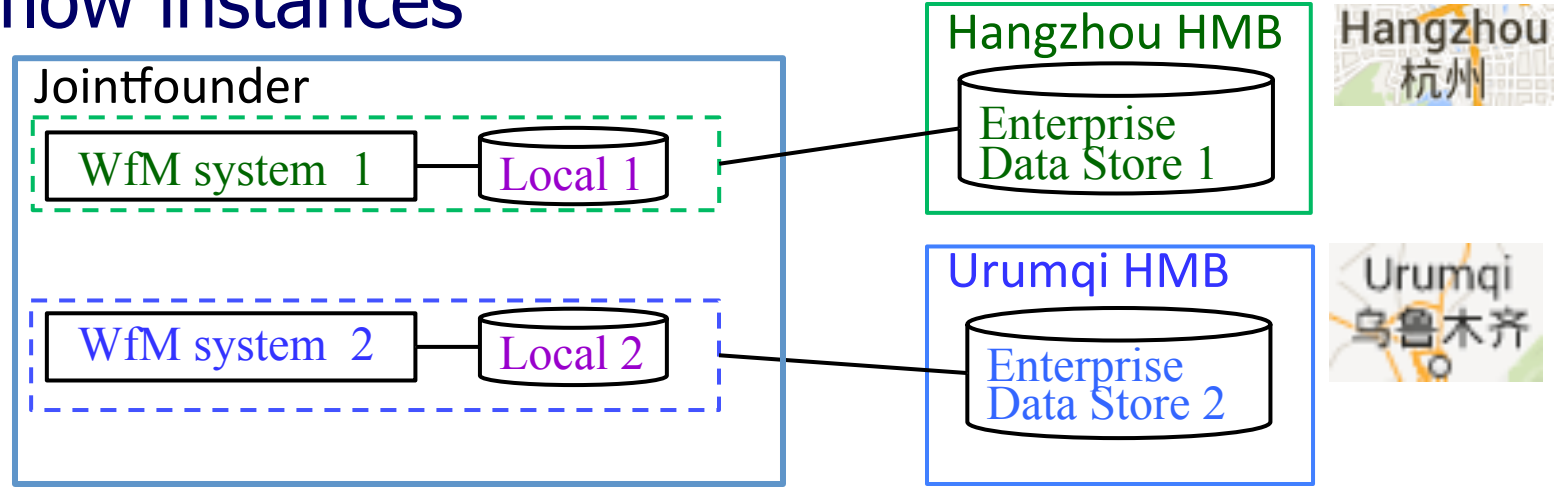
  More clients means more service technicians
  and associated management costs

Can Cloud be a Solution?

- Ideal:



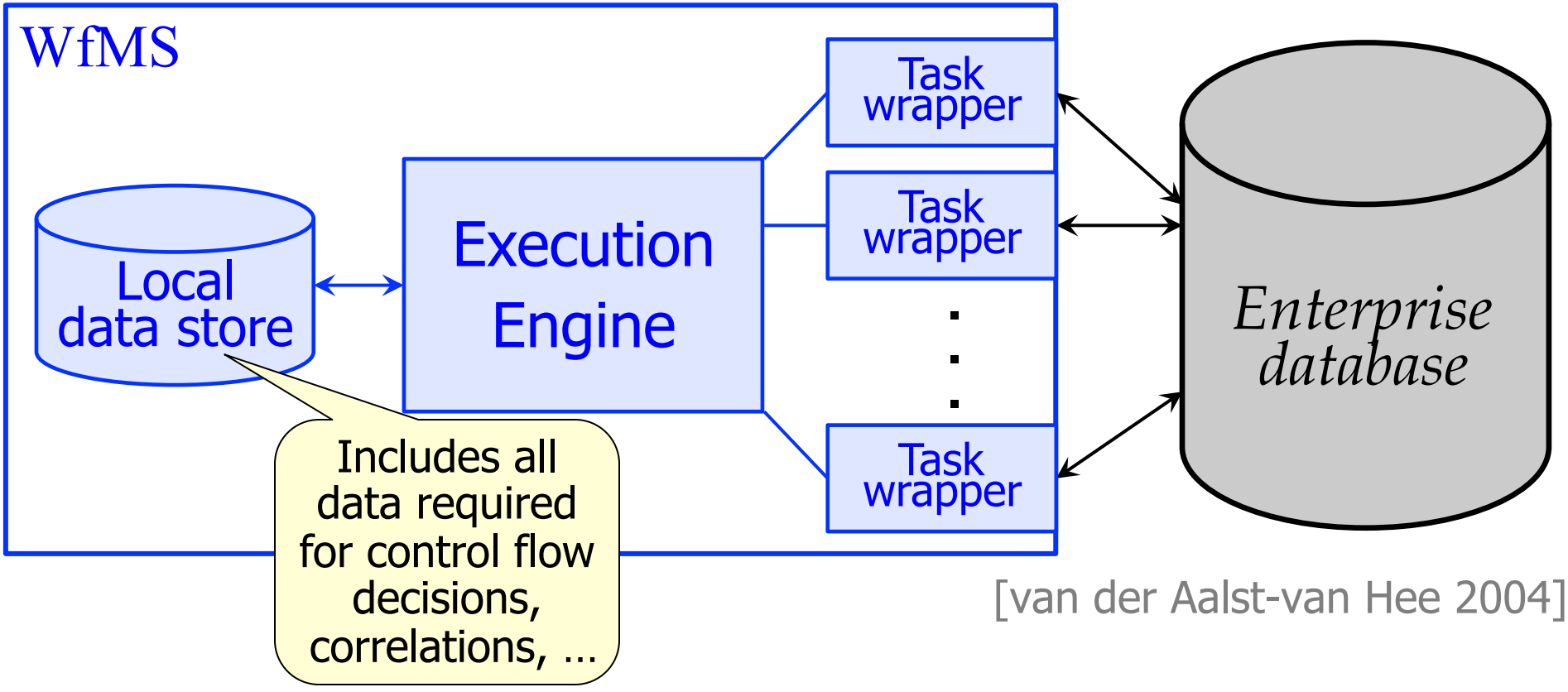
- But only naïve approach: Run one WfM system for each client HMB due to disparate local data for each HMB's workflow instances



- Shifts but does not reduce effort/cost in addressing clients technical problems: failures, changes, analytics, ...

Data Management in Workflow Systems

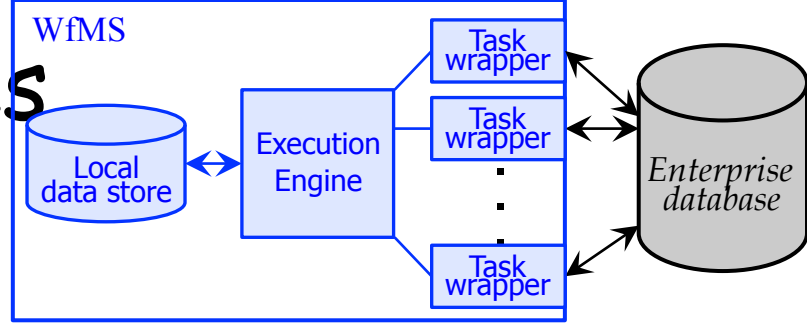
- Typical architecture:



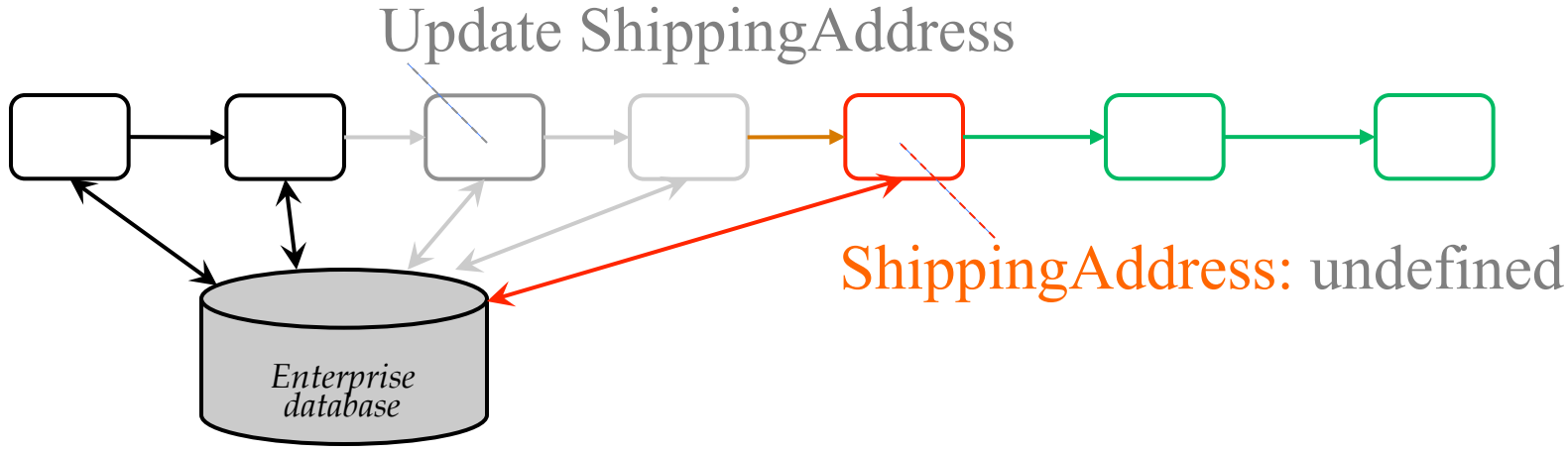
[van der Aalst-van Hee 2004]

- During execution, data can be held in each of the shaded boxes (shapes)
  - Problems? Plenty!

Example Scenario: Failures



- Enterprise database fails
  - DBMS does recovery, but data may not be consistent with data in the local store, engine, and wrappers



- Similar: Local data store fails
  - Again, recovery at store, but data may not be consistent with data in EDb, engine, and wrappers

[S.-Yang EDOC-W15]

Independence of Data Management and Execution Management

Execution Independence

the freedom of changing the process execution system while leaving conceptual BP models unchanged and vice versa

- Clean separation of responsibilities
  - WfMS: Execution
  - DBMS: Data
- Allows *Divide-and-Conquer* for management functions
  - Helps in many aspects

[Sun-S.-Yang BPM14]
[Sun-S.-Wu-Yang ICDE14]

Five Types of Data in Biz Processes

1. Specification of biz process models
2. Business data essential for business logic
   - e.g., items, shipping addresses, ...
3. Enactment status: the current execution snapshot
   - e.g., order sent, shipping request made, …
4. Resource usage and state needed for BP execution
   - e.g., cargo space reserved, truck schedule is to be determined, …
5. Correlation between processes instances
   - e.g., 3 warehouse fulfillment process instances for Jane's order (instance), …

- Traditional biz process models are weak in modeling data (types 2-5)

[Sun-S.-Yang BPM14]
[Sun-S.-Wu-Yang ICDE14]

Universal Artifacts (UA)

A universal artifact contains everything an engine needs

- A traditional business artifact:

  (*Entity information model*, *Entity lifecycle model*)

- A universal artifact contains everything an engine would need:

  (BP specification, *Entity*, States, Dependencies, *L*)

A document with the specification of the *entity lifecycle* model

actual business data

current states

correlations resources

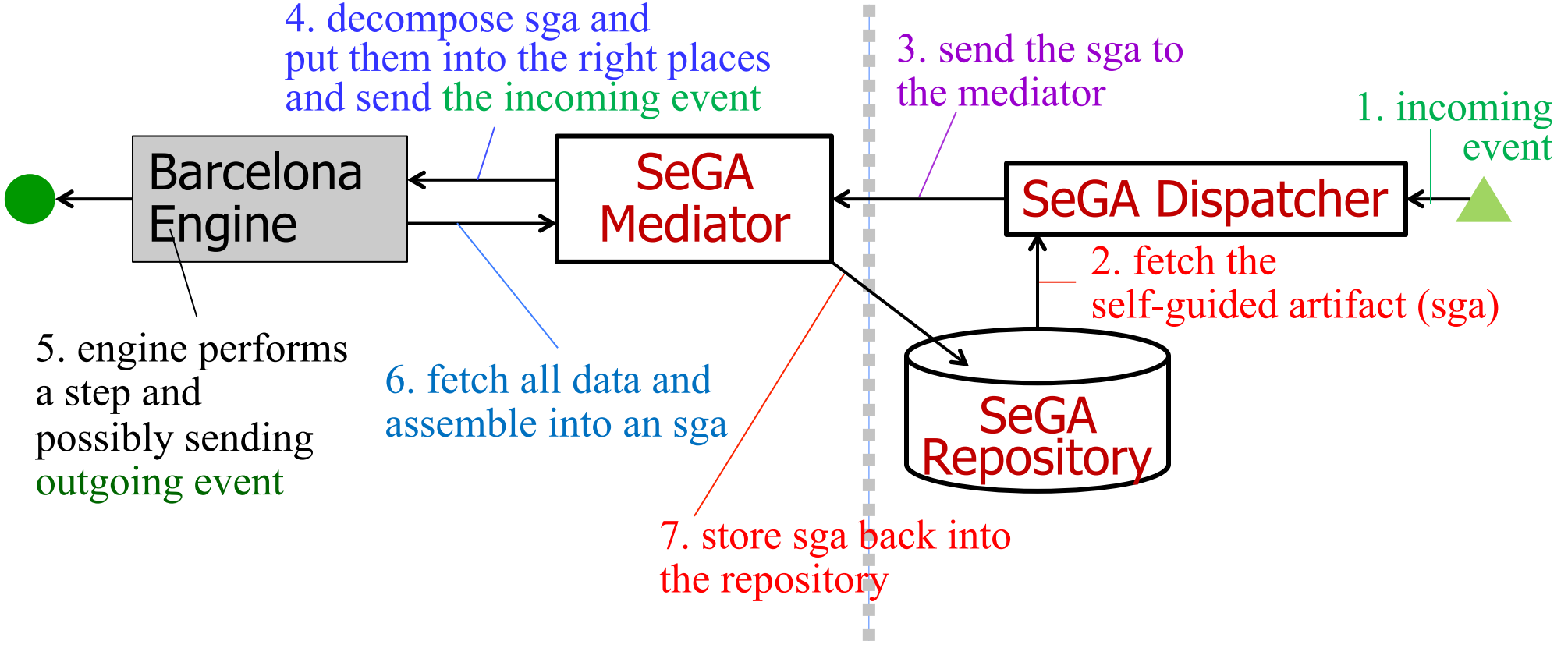
ID of modeling language

[Sun-S.-Yang BPM14][Sun-S.-Yang TMIS16]

The SeGA Framework

[Sun-S.-Yang BPM14]

- Key idea: a process wrapper to supply all data (i.e., "universal" artifact) when the engine needs to run

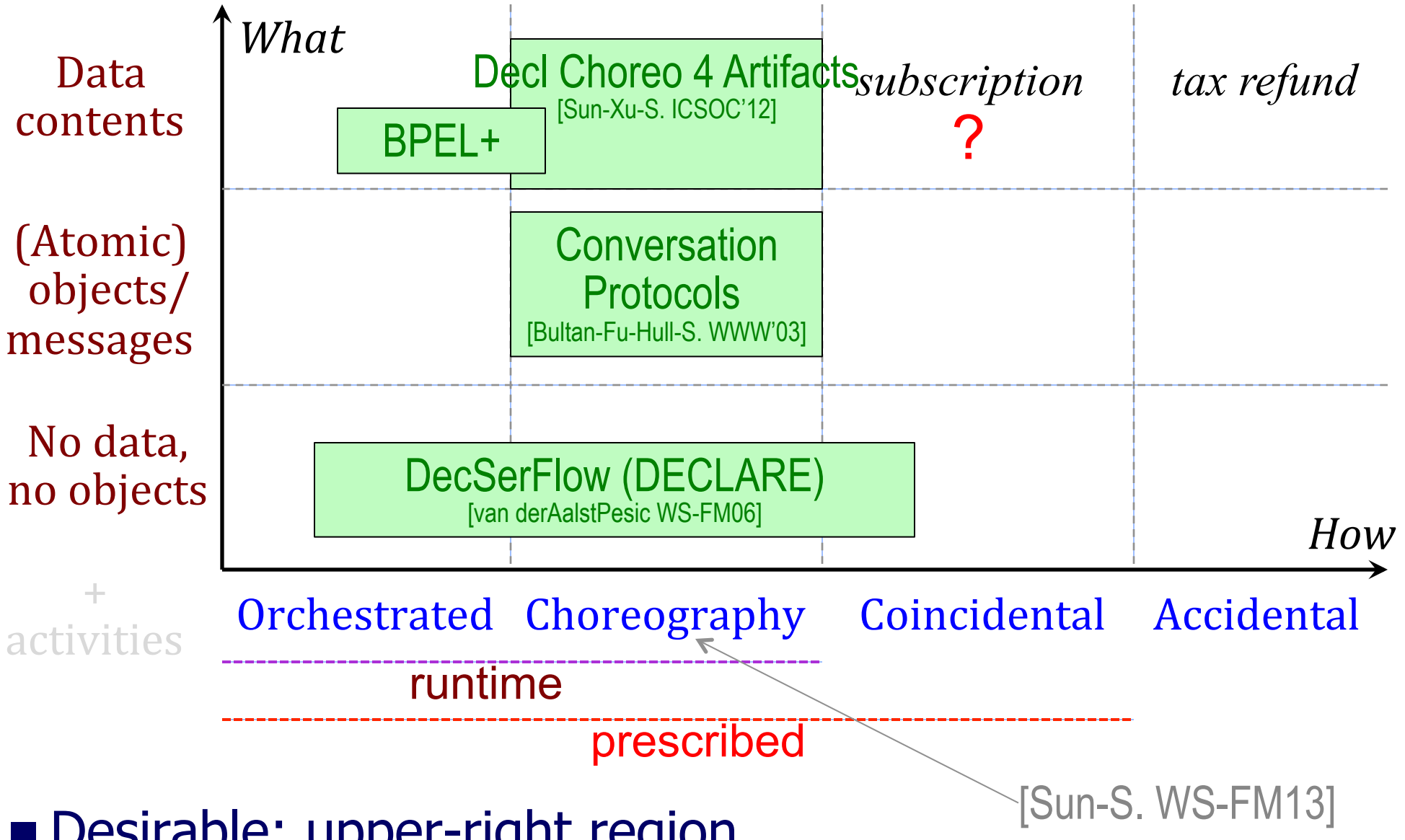


- Both Barcelona and EZ-Flow are integrated with SeGA
- 杭州中房信息科技有限公司 Hangzhou Jointfounder InformationTechnology Co., Ltd : prototype (RMB1.2M, 2014-16)

Plan for the Talk

- The Need for Process-Process Relationships
- Process Design and Modeling with Data
- Runtime Management
- Towards Process-Relationship (PR) Modeling
- Further Challenges

(Modeling) Process Relationships

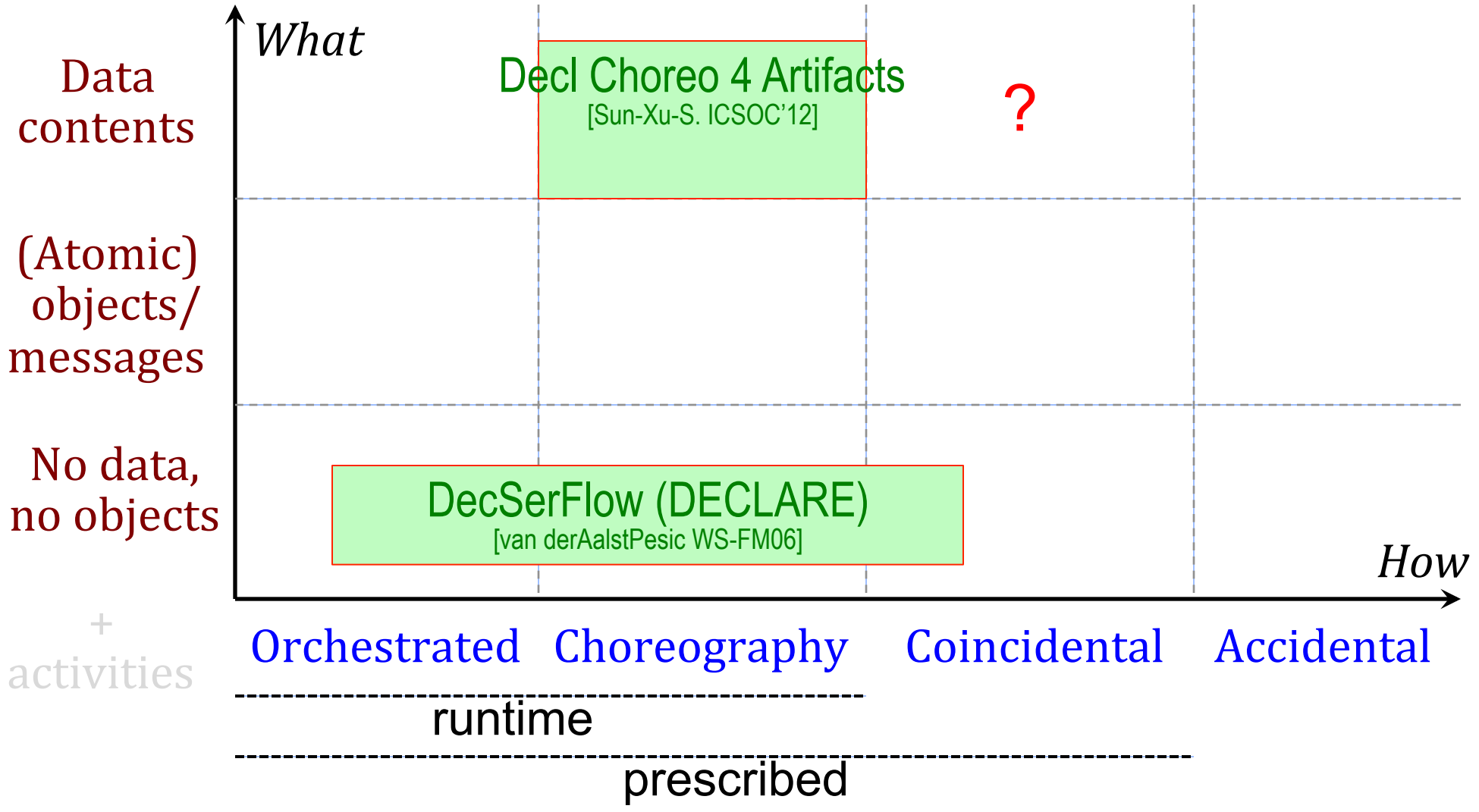


- Desirable: upper-right region

Three Types of Process Relationships

- Consider binary relationships
- Occurrences: how their instances should be related
  - e.g., *adding a driver to an auto policy causes new insurance cards to be sent*
- Cardinality: How many instances should be related
  - e.g., *if a posted charge reaches 80% of the credit limit, at most 3 warning messages should be sent at a 3-day interval*
- Data: the relationship depends on the data content
  - e.g., *if the reimbursement total exceeds CA$5000, dean's approval is necessary*

(Modeling) Process Relationships

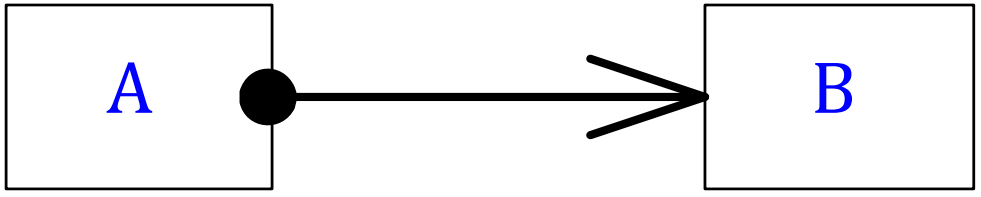


- Desirable: upper-right region

DecSerFlow (DECLARE)

[van derAalst-Pesic WS-FM06]

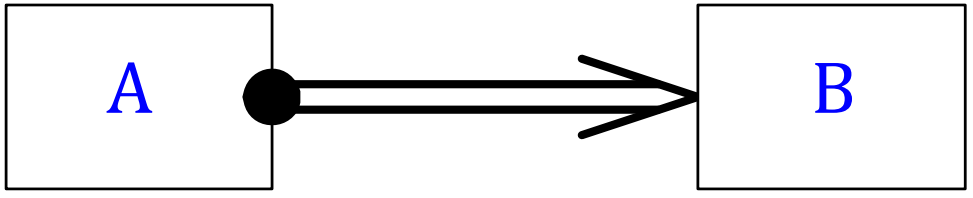
- Modeling language constraining behaviors (executions) through occurrences of activity instances
- Unary: number of executions of an activity
- Binary: (co-)existence, response, precedence, etc.
  - Alternative, succession
- N-ary constraints are possible, negation is also allowed
- Example: Every A is followed by a B (response)



DecSerFlow (DECLARE)

[van derAalst-Pesic WS-FM06]

- Modeling language constraining behaviors (executions) through occurrences of activity instances
- Unary: number of executions of an activity
- Binary: (co-)existence, response, precedence, etc.
  - Alternative, succession
- N-ary constraints are possible, negation is also allowed
- Example: Every A is followed by a distinct B (alt. resp.)

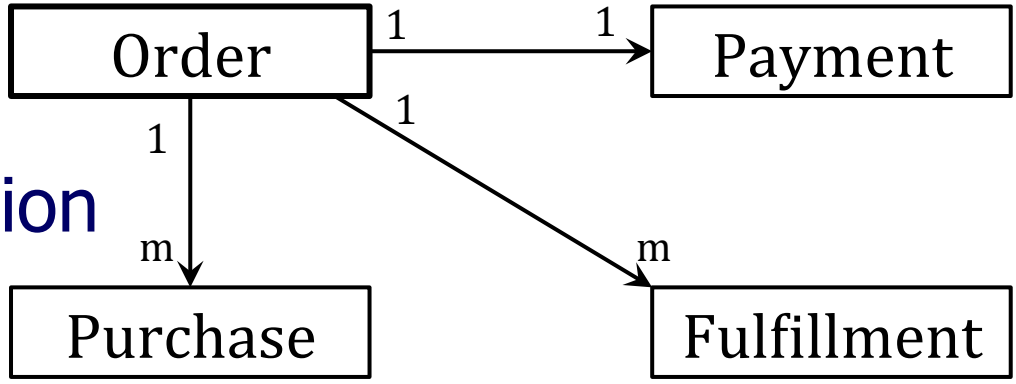


- Needed: cardinality, data

Correlation Diagram

[Sun-Xu-S. ICSOC12]

- Processes as rectangles
- Correlation as edges
  Direction: initiation/invocation
- Cardinality constraints
  on instances



- Choreography constraints on messages
- Examples:

$$\forall x{\in}\text{Order}\ \text{OR}(\mu, \text{ext}, x) \wedge \mu.\text{amount}{>}10\ \ (succ){\rightarrow}\ \ \text{CP}[\mu](x, \texttt{Purchase}[\mu])$$

$$\forall x{\in}\text{Fulfillment}\forall y{\in}\text{Purchase}\langle x\rangle\text{PC}(\mu, y, x) \wedge y.\text{cart.price}{>}100$$
$$(succ){\rightarrow}\ \text{RS}[\mu](\texttt{Order}\langle x\rangle, x)$$

- Needed: occurrences, cardinality

Modeling Relationships (Early Thinking)

- *Order*(ID, …, | Cost, …)

Shipping(… | …)
Cancel (|), *Invoice* (|)

$$O = \text{Order}(\text{ID}{:}\,x \mid\ ) \longrightarrow O < \exists\,\text{Shipping}(\text{OrderID}{:}\,x \mid\ )$$

$$\text{Order}(\text{ID}{:}\,x \mid\ ) \xrightarrow{<} \exists\,\text{Shipping}(\text{OrderID}{:}\,x \mid\ )$$

$$\text{Cancel}(\text{OrderID}{:}\,x \mid \text{Credit}{:}\,y) \xrightarrow{<} \exists\,\text{Invoice}(\text{OrderID}{:}\,x,\ \text{Credit}{:}\,y \mid\ )$$

- Under development [S.-Wen-Yang '17]

Enterprise Process Framework

Aiming to model biz services, 4 components:

- A data model
  - For data access by at least the biz service
- A set of processes
  - Accessing data instances of the data model
- A set of relationships between processes
  - Constraining instances of process instances
- A set of KPIs / QoSs
  - Measuring aspects of interest

Plan for the Talk

- The Need for Process-Process Relationships
- Process Design and Modeling with Data
- Runtime Management
- Towards Process-Relationship (PR) Modeling
- Further Challenges

Modeling Process Relationships

- Choice of data model: ER or XML
  - What should be included?
- Choice of process model: data is essential
  - IOPE seems natural
- Relationships:

Axiom of Anticipation for Proc. Modeling

- Combining all processes in a biz service into a single one
  - Modeling languages usually allow such
  - But not a good idea:
    -- long lasting processes
    -- harder to maintain the workflow system
- What is the right size?
- *Except for the initial event, all other events/activities should be known to happen*

Modeling Process Relationships

- Choice of data model: ER or XML
  - What should be included?
- Choice of process model: data is essential
  - IOPE seems natural
- Relationships:
  - Occurrences & cardinality
  - Data: data flow, other factors?
  - Temporal constraints
- QoSs/KPIs: realistic indicators
- Goals: enabling reasoning/analytics, and (next slides)

Optimization and Automation

- Measuring QoSs/KPIs
  - Effectiveness?
- Optimizing EPFs based on QoSs/KPIs
  - Move activities from one process to another
  - Remove redundant activities
  - Replace activities by “cheaper” version
  - Batch executions
- Automation
  - Specification of EPF to technical model?
  - Data (documents, logs, emails, …) to technical models----cognitive computing

Changes and Change Impact Analysis

- Example: add services for low income housing
  - Seems benign for existing services but could impact property tax (reduction)
- Add luxury tax:
  - Could be a fraud to avoid
- Intra-EPF impact analysis
- Inter-EPF impact analysis

- Previous work focus on individual processes

Anomaly Detection and Incident Mining

- Anomaly: a potentially undesired incident
- Tax refund example:
  - Detectable if EPFs of biz accessible
  - Algorithms to analyze EPFs of vendor with tax refund EPF of customs & border control
- Double reimbursement: similar
  - Adjunct appointments routine
  - Same reimbursement from two different institutions?

- Mining logs to discover possible anomalies?
  ISC mining may help [Winter-Rinderle-Ma EDOC17]

Conclusions

- A new approach to modeling business services
  - Cumbersome to put all in one process
    axiom of anticipation
- Process relationships – a key element in gluing together processes
  - Similar to modularity, hierarchies, yet fundamentally distinct
- Fairly green field, not much has been done
  - Techniques from data modeling might be useful