团队合作与通信

（Teamwork & Agent Communication）

毛文吉

中国科学院自动化研究所

Outline

- Understanding teamwork
- Teamwork theories
- Practical teamwork model (STEAM)
- Agent communication

Outline

- ➢ Understanding teamwork
- Teamwork theories
- Practical teamwork model (STEAM)
- Agent communication

Why Teams?

- *Teamwork* as a basis for organizing groups
  - 种姓制度
  - Not castes, master-slave, or contract relationships
- *Robust* organizations
  - Responsibility to substitute (cover for one another)
  - Mutual assistance when run into difficulties
  - Teammates ensure key information is communicated to peers
- Still capable of structure (not necessarily a *flat* organization)
  - Teams of teams
  - Different roles require different responsibilities
  - Variations in capabilities and limitations

Understanding Teamwork: Motivators

Coming together is beginning --

Keeping together is progress --

Working together is success

- *Teamwork* is the ability to work together towards a common vision
- The ability to direct individual accomplishment toward organizational objectives
- It is the fuel that allows common people to attain uncommon results

Understanding Teamwork 共同目标

协同动作

- Not just a union of simultaneous coordinated actions
- Different from contracting

- Ordinary traffic ✗ 成员不知道在team中
- Driving in a convoy（车队）
  - Two friends *A* & *B* together drive in a convoy ✓
  - Person *B* ’s car is secretly following *A* ’s car ✗
- Pass play（传球）in Soccer ✓
- Table tennis doubles（双打）✓
- Contracting with a software company ✗ 不是团队成员
- Marathon ✗

  乐队演奏
- Orchestra performance ✓

Understanding Teamwork

- Not just a union of simultaneous coordinated actions
- Different from contracting

- Ordinary traffic
- Driving in a convoy
  - *Two friends A & B together drive in a convoy*
  - *B is secretly following A*
- Pass play in Soccer
- Table tennis doubles（双打）
- Contracting with a software company
- Marathon
- Orchestra performance

What is Teamwork?

Cooperative effort by members of a team to achieve a common goal

- Common goal, teammate cooperate, help each other, *team spirit?*

Research questions:

- Build *robust* teams with *flexible coordination* in *complex* domains
- Team *formation*
- Teamwork *monitoring*, diagnosis, analysis, explanation
- Collaborative *negotiation*
- *Adjustable* autonomy in teams 自适应
- Team plan *recognition*
- *Learning* in teamwork ...

How to Build Effective Agent Teams?

领域级别

Given *multiple agents* with domain-level skills & team goals

- Appropriate teamwork/coordination to achieve team goals?
  - *When/with whom/what to communicate? Take up others' tasks?*

Difficulties:

- Members' dynamically obtain differing, incomplete information
- Communication time-consuming, costly, risky
- Members unexpectedly fail in fulfilling tasks
- Uncertainty about what other team members know

Outline

- Understanding teamwork
- Teamwork theories
- Practical teamwork model (STEAM)
- Agent communication

Teamwork Theories

- Theory: *Fundamentally* understand teamwork
  - Is there one underlying model for working in teams?

    领域知识

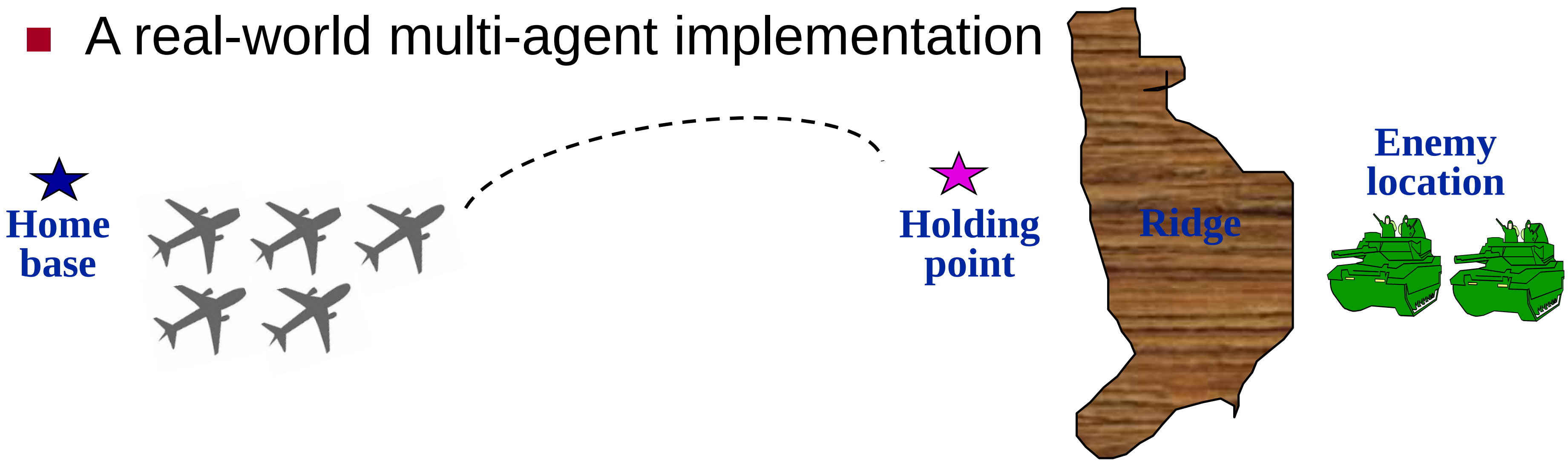
    e.g. is the teamwork in Soccer the same as in other domains?
  - Are there some fundamentally different types of teamwork?

- Practical benefits: *Specification* for teamwork applications
  - Build reusable teamwork code 建立复用的 team code
  - Teamwork libraries: High-level team-oriented programs

    高级别 面向团队编程
  - Build robust applications

Ad Hoc Team Coordination in Real World

- A real-world multi-agent implementation



- Execute-Mission
  - Fly-flight-plan
    - Fly-route
      - Low-level
      - Noe (Notice of exception)
      - …
    - …
  - Engage
    - Employ-weapons
      - Mask
      - Unmask
      - Employ-missile

- **Pilot plan hierarchy**
- **Ad-hoc team coordination**

临时的

Problems with Ad Hoc Team Coordination

- For *2-3* agents, carefully *controlled* coordination may work
  - Individual agents provide individual goals and plans
  - Based on fixed, domain-specific coordination plans
  - *No framework* to anticipate failures; numerous ad hoc plans
- ➢ Inflexibility, unexpected failures, lack of reuse
- For *bigger* teams and *complex* scenarios, ad hoc coordination can cause serious *problems*:
  - Commander returned to home base alone, leaving others in battlefield
  - Company waited indefinitely for message, after the scout crashed
  - Flight leader got order first and started mission alone, others not ready
  - Scout never returned, while others got into infinite loop …
- ➢ **Need the support of teamwork theories**

Different BDI Theories of Teamwork

Explicitly based on commitments, beliefs, desires, intentions:

- *Joint Intentions* [Cohen & Levesque 90]
- *SharedPlans* [Grosz & Kraus 96]
- *Planned team activity* [Sonenberg et al 92]
- *We-intentions* [Searle 90]

- Basis for effective team building
- But did not cover *team formation*, *negotiation*, *learning*, ….

Focus on Joint Intentions Theory

Commitment is central to *teamwork*:

- *Team* commitment rather than *individual* commitment

Next we shall

- Go through the main theory *step by step*
- Help understand why the theory is the *way* it is
- Also help understand some of the *issues* in teams

Recall: Commitments Fundamental to Agents

Agents plan because they are *resource bounded* (Recall Bratman)

- Agents must form and **commit** to plans

  Commit == *hold on to a plan, don't drop it easily*

- Commitments constrain reasoning:
  - Frame problems for an agent (*what to reason about*)

    Agent needs to determine how to fulfill commitments
  - Provide filter of admissibility (*what not to reason about*)

    Agent filters out options incompatible with commitments

- *Intention* is kind of *committed* plan (e.g. my next trip)

Formalizing Commitments in Modal Logic

- Notation (Cohen & Levesque 1990):
  - Bel (x P): Agent x has P as a belief
  - Goal (x P): Agent x has P as a goal
  - (Eventually P): Somewhere in the future P becomes True
  - (Until P Q)/(Q Until P): Q is True until P True
  - (Always P): (NOT (Eventually NOT(P)))
  - (Never P): (Always (NOT(P))
  - AND, OR, NOT: Logic connectives

Recall: Individual Commitments

Focus on achievement goals:

- If P is a *commitment*, then don't drop P as a goal easily

  (Goal x (Eventually P)) UNTIL

  - P Achieved: (Bel x P) OR
  - P Unachievable: (Bel x (Never P)) OR
  - P Irrelevant (Q is possibly the *reason* for P): (Bel x NOT(Q))

Commitments as persistent goals (*PGOAL*, 持续的目标):

- (PGOAL x P Q) is defined as:

  (Bel x NOT(P)) AND (Goal x (Eventually P)) AND

  (UNTIL ((Bel x P) OR

  (Bel x (Never P)) OR

  (Bel x NOT(Q)))

  (Goal x (Eventually P)))

Formalizing Team Commitments: Attempt #1

Attempt #1: Sum of individual commitments

- Team: T={x, y} (Two individuals)
- Team task: searching for a lost object: SFLO
- P=(Done T SFLO); Q=High-level cause for P

Team commitment:

(PGOAL x *X's-part*(P) Q) AND (PGOAL y *Y's-part*(P) Q)

Problem:

- X and Y do *not even know* they are a team
- Just *individual* simultaneous effort

Defining Mutual Beliefs（共同的信念）

- Mutual Beliefs (MB T P)

  (MB x y P):
  - (Bel x P)
  - (Bel y P)
  - (Bel x (Bel y P))
  - (Bel y (Bel x P))
  - (Bel x (Bel y (Bel x P)))
  - (Bel y (Bel x (Bel y P)))
  - … …
- Infinite levels of nesting in principle
- Practical systems find compromises
- Remains an active area of research

Formalizing Team Commitments: Attempt #2

Attempt #2: Add MB to individual commitments

- Team: T={x, y} (Two individuals)
- Team task: searching for a lost object: SFLO
- P=(Done T SFLO); Q=High-level cause for P

Team joint commitment:

(Joint-PGOAL T P Q
(MB T (PGOAL x *X's-part*(P) Q) AND
(PGOAL y *Y's-part*(P) Q)))

Problem:

- One team member can *drop* P and leave, e.g. if (Bel x P)
- *Glue* to hold team together is missing

We-Intentions (Searle 90)

- *Collective* intentional behavior is a *primitive* phenomenon
- *Cannot* be analyzed in terms of *individual* intentions
  - Puppies playing in a lawn, practicing pass play …
  - No single individual acting alone or accomplishing results alone

- "*We*" are doing actions "*together*": collective intention
- Since collective intention is *irreducible*, how to formalize it ?

Formalizing Team Commitments: Attempt #3

Attempt #3: Team as a single agent

- Team: T={x, y} (Two individuals)
- Team task: searching for a lost object: SFLO
- P=(Done T SFLO); Q=High-level cause for P

Team joint commitment:

(Joint-PGOAL T P Q

(MB T NOT(P)) AND (MG T (Eventually P)) AND

(UNTIL [(MB T P) OR

(MB T (Never P)) OR

(MB T NOT(Q))]

(MG T (Eventually P))))

- Mutual Goal (MG T P) is: (MB T ((Goal x P) and (Goal y P)))

Problems in Attempt #3: Too Strong

Team is not an individual and hence attempt fails!

- As time goes by, individuals may *diverge* in their beliefs (though initially beliefs were synchronized)
- E.g. after some passage of time, x *privately* comes to believe that P is impossible, but y does not believe so yet

This divergence in private beliefs is not allowed.

<u>Example</u>: *Violate* our current definition of JPG

- Suppose (Bel x (Never P)), then x *drops* (Goal x (Eventually P))
  - Otherwise, beliefs and goals are inconsistent
- So (MG T (Eventually P)) is *false*, but *UNTIL* condition is *false*

(As there is no mutual belief formed yet, this MG is supposed to persist until there is a certain mutual belief)

Formalizing Team Commitments: JPG

- Joint Persistent Goal (JPG, 联合持续的目标)

Team joint commitment:

(JPG T P Q

(MB T NOT(P)) AND (MG T (Eventually P)) AND

(UNTIL [(MB T P) OR

(MB T (Never P)) OR

(MB T NOT(Q))]

(WMG T (Eventually P) Q)))

- Team *mutually* believes the goal P is not achieved
- Team has the *mutual goal* to achieve P
- Until Team *mutually* believes P *achieved, unachievable or irrelevant:*
  - Team has P as a *Weak Mutual Goal* (WMG, 弱的共同目标)

Formalizing Team Commitments: JPG

- Joint Persistent Goal (JPG, 联合持续的目标)

Team joint commitment:

(JPG T P Q
  (MB T NOT(P)) AND (MG T (Eventually P)) AND
  (UNTIL [(MB T P) OR
    (MB T (Never P)) OR
    (MB T NOT(Q))]
  (WMG T (Eventually P) Q)))

- Weak Mutual Goal (WMG T P Q) is defined as:
  (MB T (WG x T P Q) and (WG y T P Q))
- Team *mutually* believes that each member has P as a *Weak Goal*

Formalizing Team Commitments: WG

- Joint Persistent Goal (JPG, 联合持续的目标)

Team joint commitment:

(JPG T P Q

(MB T NOT(P)) AND (MG T (Eventually P)) AND

(UNTIL [(MB T P) OR

(MB T (Never P)) OR

(MB T NOT(Q))]

(MB T (WG x T (Eventually P) Q) and (WG y T (Eventually P) Q))))

- Team *mutually* believes the goal P is not achieved
- Team has the *mutual goal* to achieve P
- Until Team *mutually* believes P *achieved, unachievable or irrelevant:*
  - Team *mutually* believes that each member has P as a *Weak Goal*

Weak Goals (WG, 弱的目标)

- UNTIL clause only contains weak goal:

(WG x T (Eventually P) Q) is defined as:

[(Bel x NOT(P)) AND (Goal x (Eventually P))]

OR

[(Bel x P) AND (Goal x (Eventually (MB T P)))]

OR

[(Bel x (Never P)) AND (Goal x (Eventually (MB T (Never P))))]

OR

[(Bel x NOT(Q)) AND (Goal x (Eventually (MB T NOT(Q))))]

Example: Divergence of Private Beliefs

Divergence of private beliefs is okay now.

Example: Initially team has JPG, but now (Bel x (Never P))

- If (Bel x (Never P)), then x *drops* (Goal x (Eventually P))
- We enter the *UNTIL* clause, but with a change:

(UNTIL [(MB T P) OR
(MB T (Never P)) OR
(MB T NOT(Q))]
(MB T [(Bel x (Never P)) AND (Goal x (Eventually (MB T (Never P))))]
AND (WG y T (Eventually P) Q)))

- So x must commit to attain MB of (Never P) (e.g. communicate y)
- Cannot simply drop P and walk away, must inform others

JPG Exercise

- Suppose T consists of one individual x, then we have

  (JPG x x P Q) = (PGOAL x P Q)

- Note that

  (MB x x P) => (Bel x P)

  (MB x x Never(P)) => (Bel x Never(P))

  (MB x x NOT(P)) => (Bel x NOT(P)) => NOT(Bel x P)

  (MG x x (Eventually P)) => (Goal x (Eventually P))

  (WMG x x P Q) => (WG x x P Q)

JPG and WG: Implications

- Joint Persistent Goal (*JPG*) as team *joint commitment*
  - Team members must hold this mental attitude
  - bind the team together
- *Weak mutual goal*: team's *MB* of each member's *weak goal*
  - Team members will attempt to achieve P
  - If x privately believes P is *achieved*, *unachievable* or *irrelevant*
    - *Commit to make P mutually believed*
- Implications:
  - *Jointly commit* to a team goal and work together *persistently*
  - Also accommodate to the *divergence* of individual beliefs, but keep teammates *informed*
    - Lead to communication in teamwork

Joint Intentions

- Joint commitment basis of *joint intention*:
  - Joint commitment to a joint activity while mutually believing that agents are about to do the joint action
  - Must deliberately do it (jointly)

Definition 5 (JI x y a q) $\stackrel{\text{def}}{=}$

```
(JPG x y
     (DONE x y
              [UNTIL (DONE x y a)
                     (MB x y (DOING x y a))]?;a)
     q)
```

P

  - W?;a → Action a with W holding initially
  - (Done T a) → Similar to P

SharedPlans (Grosz & Kraus 96)

- No joint mental attitude, instead "*intend-that*" for helpful behavior
- *SharedPlan* of a group GR requires that:
  - Mutual belief that each member intends that GR achieves *joint goal*
  - Mutual belief (agreement) of the *joint recipe*
  - For each step in the *recipe*:
    - Some individual/subteam forms *SharedPlan* for that step
    - Other members *intend that* individual/subteam perform the step
- *SharedPlans* may be *partial* (e.g., recipe not fully elaborated and can be developed *incrementally*)

Example: SharedPlan

- “Intend-that” is defined via several axioms
- E.g. *Axiom A7* in *SharedPlan*:
  - Group GR has a *sharedplan* S1
  - G1 is a member of GR, G2 is a member of GR
  - G1 can perform A1
  - G1 intends that for G2 to bring about some action A2 in service of S1
  - G1 believes that G1’s performing A1 and then G2’s performing A2 will be cheaper (lower cost) than G1 not performing A1 and G2 performing A2
- Then G1 will consider performing A1

Exercise: Quick Questions

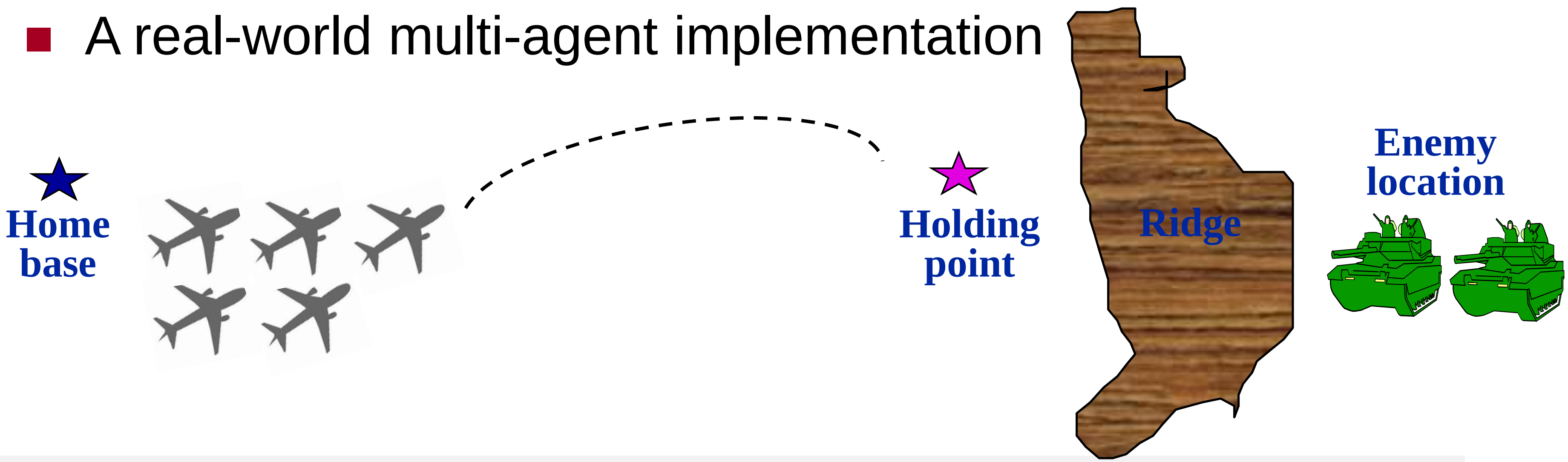
- In Cohen & Levesque's joint intentions theory, is joint commitments a primitive concept? Yes/No NO
  最基本的概念
- If No, then what are the primitive concepts from which joint intentions theory is built?
- Does Joint Intentions theory cover all of teamwork? Yes/No NO
- What are some key problems in the joint intentions theory? 失败．团队扩展．团队生成

Outline

- Understanding teamwork
- Teamwork theories
- Practical teamwork model (STEAM)
- Agent communication

Ad Hoc Team Coordination in Real World

- A real-world multi-agent implementation



- Execute-Mission
  - Fly-flight-plan
    - Fly-route
      - Low-level
      - Noe (Notice of exception)
      - …
    - …
  - Engage
    - Employ-weapons
      - Mask
      - Unmask
      - Employ-missile

- **Pilot plan hierarchy**
- **Ad-hoc team coordination**

Avoiding Ad Hoc Team Coordination

- For *2-3* agents, carefully *controlled* coordination may work
  - Individual agents provide individual goals and plans
  - Based on fixed, domain-specific coordination plans
  - *No framework* to anticipate failures; numerous ad hoc plans

➢ Inflexibility, unexpected failures, lack of reuse

- For *bigger* teams and *complex* scenarios, ad hoc coordination can cause serious *problems*:
  - Commander returned to home base alone, leaving others in battlefield
  - Company waited indefinitely for message, after the scout crashed
  - Flight leader got order first and started mission alone, others not ready
  - Scout never returned, while others got into infinite loop …

➢ **Need the support of teamwork models**

Explicit General Teamwork Models

- Provide agents with *explicit*, *general model* of teamwork:
  - Outline team members' commitments, responsibilities
  - General team coordination algorithms

Advantages:

- *Robustness* in the presence of unanticipated events
  - Agents themselves plan responses to teamwork contingencies
- *Reuse/transfer* to other applications:
  - Why build coordination from scratch each time?

Teamwork models:

- GRATE* [Jennings 95]
- COLLAGEN [Rich & Sindner 97]
- STEAM [Tambe 97], Java TEAMCORE [Scerri *et al* 03] …

Relationships to Teamwork Theories

Based on teamwork theories: *Joint Intentions*, *SharedPlans*

- Joint Intentions
  - Agents behave responsibly towards teammates
  - Once jointly committed to a team goal, inform teammates if it privately believes team goal terminated
- SharedPlans
  - Evolving hierarchy of commitments; not just one goal
  - "Intend-that" to fulfill unfulfilled tasks
  - Individuals may volunteer or request others to take up tasks

➢ *Joint Intentions* provide commitments; *SharedPlans* are suited in complex tasks

STEAM: A Practical Teamwork Model

STEAM (Tambe 97): A Shell for Teamwork

Key points:

- *Operationalize* teamwork theories in agent architectures
  - Modal-logic specifications operationalized in tractable algorithm
- Apply and reuse teamwork model in *real-world* domains
  - Empirically illustrate cross-domain general teamwork principles
- Agents automatically *coordinate* team communication/retasking
  - Enable team-oriented programming
- Address some major issues:
  - Communication costs, uncertainty; Team member failure, …

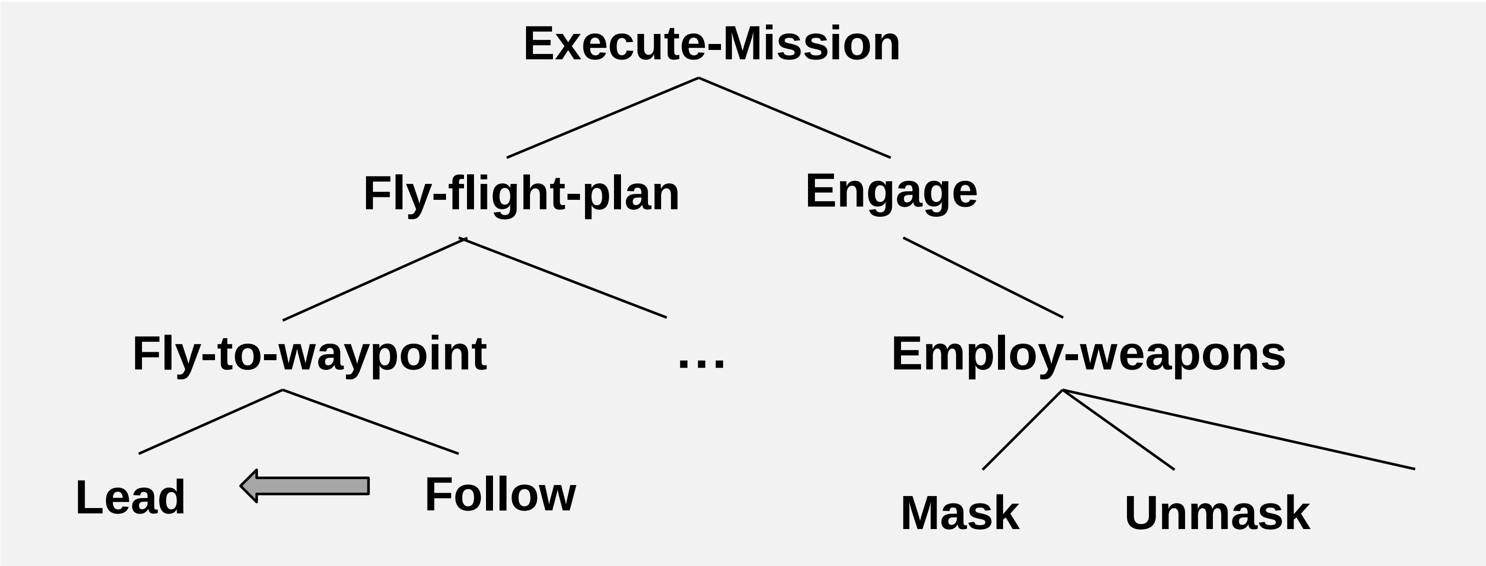
M. Tambe. Reactive Towards Flexible Teamwork. *Journal of AI Research*, 1997.

STEAM: Plans and Constraints

Explicitly using situated/reactive team plans:

- Hierarchically expand into individual/subteam plans
  - Each plan has preconditions, body, termination conditions
- Coordinate constraints among plans (e.g. Follow→Lead)

Hierarchical situated plans:

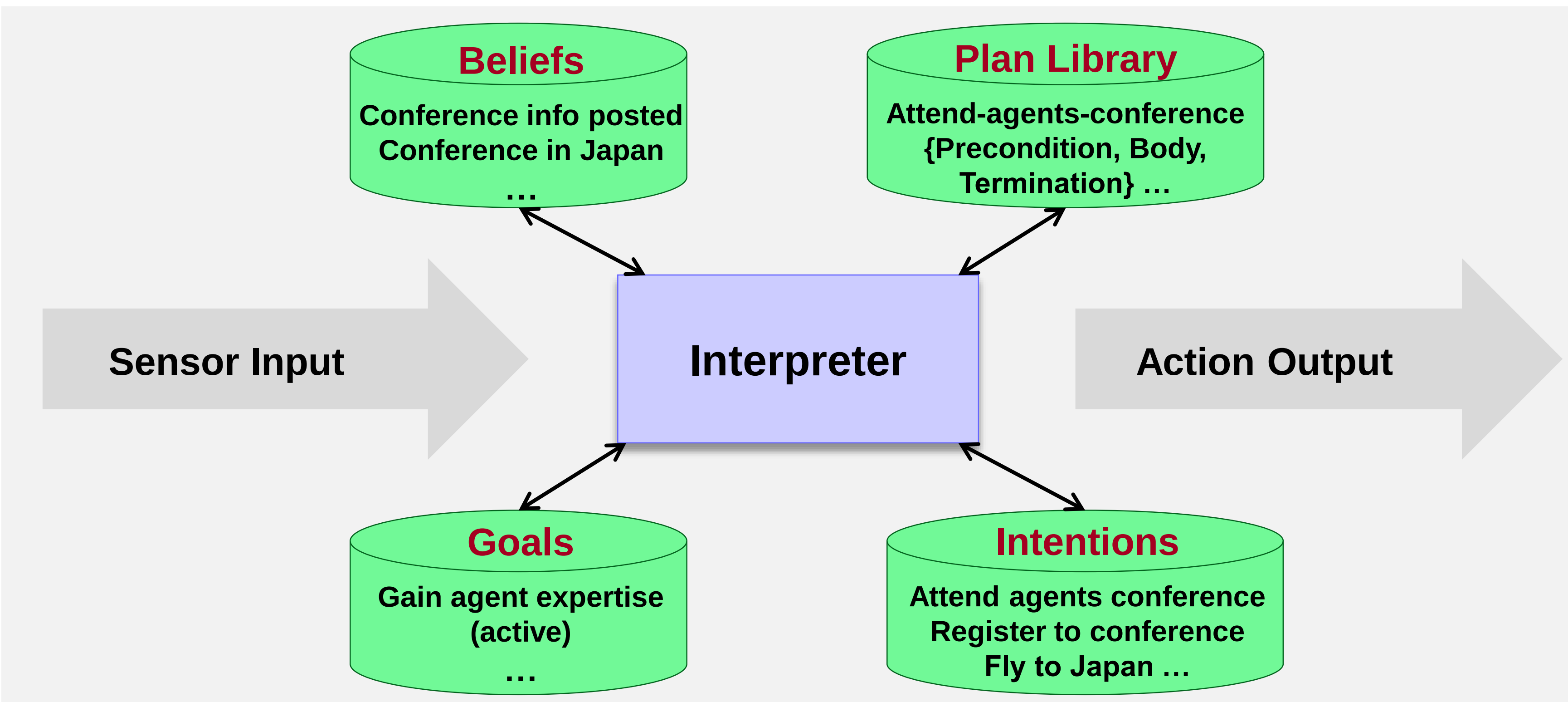


Recall: Situated/Reactive Plans（情境规划）

- Practical implementations may use *situated plans*
- Situated/Reactive plans consist of
  - *Preconditions*: matched with agents' beliefs to activate plan
  - *Termination conditions*: to terminate plan when matched
  - *Plan body:* to execute when plan is activated
- Example: Plan *Attend-agents-conference*
  - Precondition: *Call for participation posted*
  - Body: *Register, fly to site, attend sessions, fly back…*
  - Termination condition*: Agents conference attended*
- Basis of several different types of agent architectures
  - PRS, SOAR, IRMA, InteRRaP, JACK™ …

Recall: PRS-Type BDI Architecture

- Integrate BDI model, planning and reactive techniques



Hierarchical Situated Plans

- Evacuation (Rescue) domain

Team plan:

Fly-flight-plan

Precondition: Flight plan
Termination: Reach destination
See enemy
Plan body: (Null)

Subteam plan:

Fly-route   Fly-to-waypoint

Precondition: Next waypoint
Termination: Reach waypoint
Plan body: Search for lead helo

Individual plan:

Fly-lead ⇐ Follow-leader

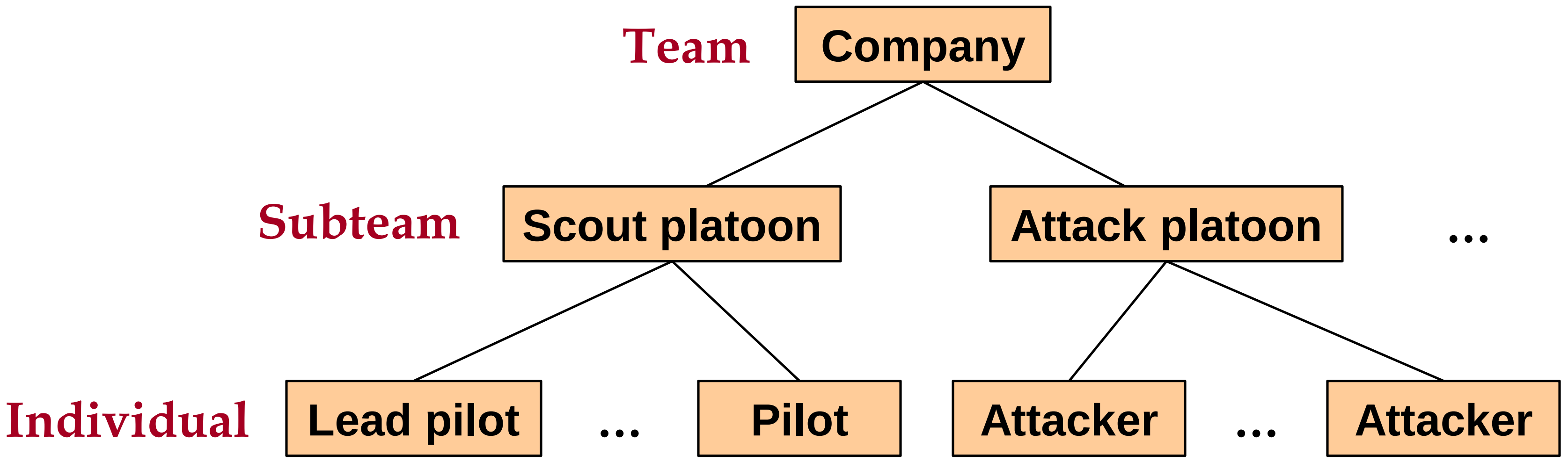
Precondition: See lead helo
Termination: Reach waypoint
Plan body: Heading/speed matches lead

STEAM: Organization Hierarchy

Specifying hierarchical organization structure

- Organization hierarchy of team, subteam and individual with different roles

Organization hierarchy for agents:

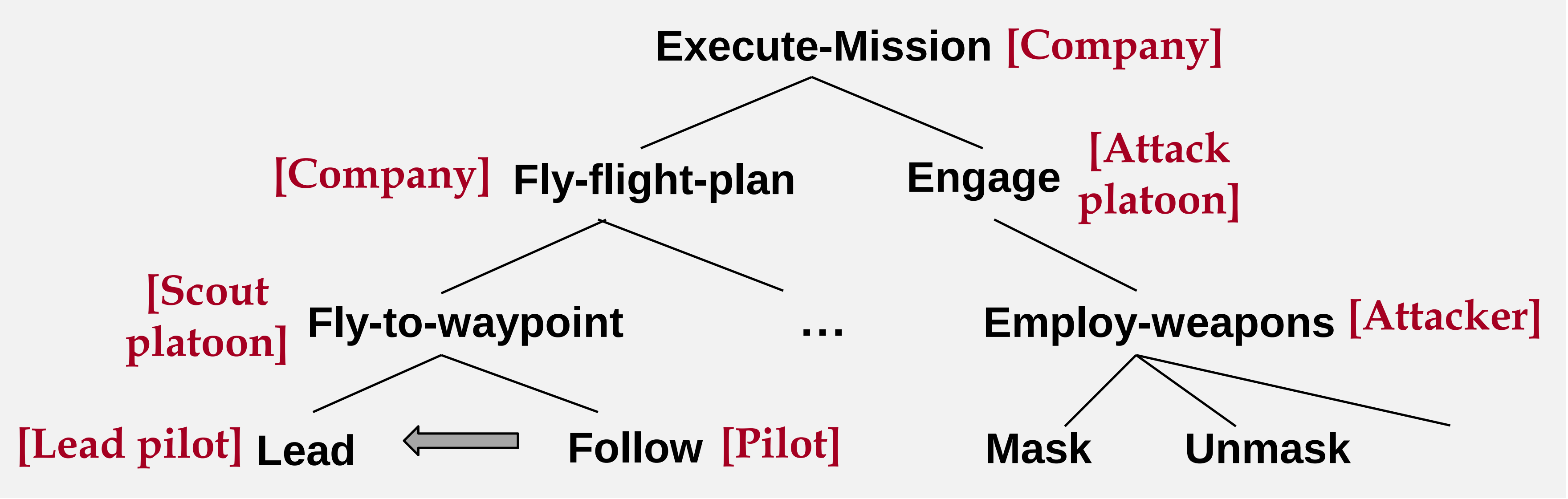


STEAM: Role Assignments

Using situated team plans for role assignments

- Assign roles in the organization hierarchy to multi-agent team plans

Role assignments in evacuation plan:



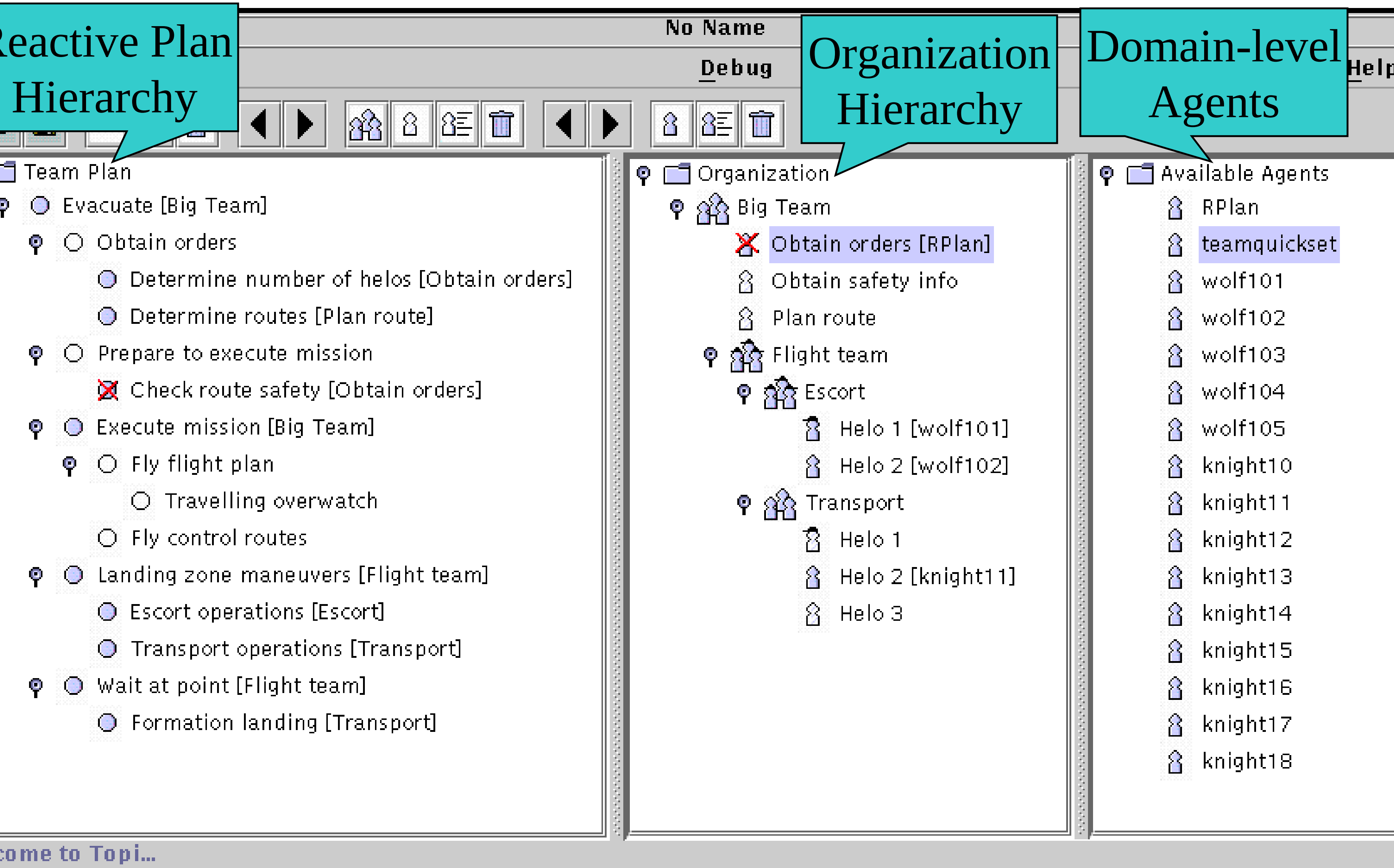
STEAM: Team-Oriented Programs

STEAM realizes team-oriented programming with

- Reactive team plan hierarchy
- Organization hierarchy of roles for individuals and teams
- Assignments of roles to plans in the team hierarchy
- Finally,

  Assignments of agents to roles specified in the organization hierarchy

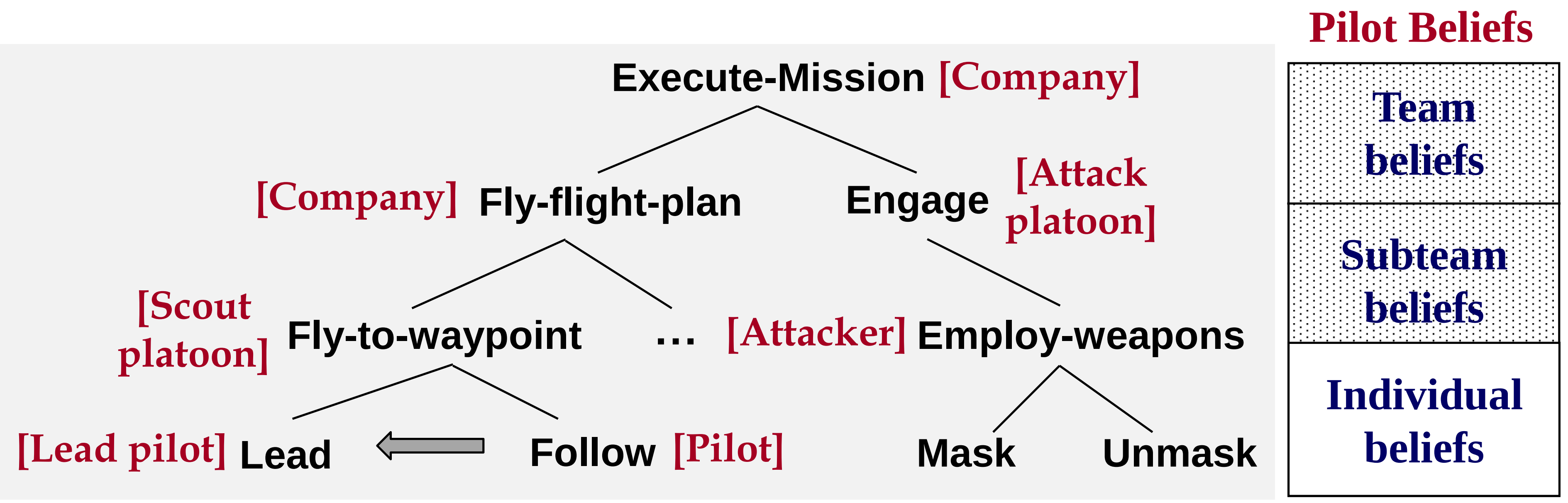
Team-Oriented Program Interface



STEAM: Team Members at Run-Time

At run-time, each agent maintains a copy of team-oriented program

- Beliefs: Team beliefs, subteam beliefs, individual beliefs
  - No Shared memory
- Individual plans modify individual beliefs, but not team beliefs



STEAM: Team Plan Execution

- All team plans are executed by *forming* and *terminating* joint commitments:
  - Request-confirm exchanges so all team members *select* & *commit*
  - Establish MB for *achieved/unachievable/irrelevant* (*AUI*) to terminate
  - Forming and terminating team plans: All *communication* in STEAM

Example: Team of company jointly commit to "*execute mission*" plan

- If commander privately believes *mission achieved* …
- Commander must establish mutual belief in termination condition
- It communicates *mission achieved*, so no one is left behind

- Hierarchy of jointly committed team plans & subteam plans
  - Team is coherent when executing and terminating team plans

Hierarchical Situated Plans

- Evacuation (Rescue) domain

Team plan:

Fly-flight-plan

Precondition: Flight plan
Termination: Reach destination
See enemy
Plan body: (Null)

Subteam plan:

Fly-route   Fly-to-waypoint

Precondition: Next waypoint
Termination: Reach waypoint
Plan body: Search for lead helo

Individual plan:

Fly-lead ⇐ Follow-leader

Precondition: See lead helo
Termination: Reach waypoint
Plan body: Heading/speed matches lead

STEAM: Monitor and Repair

Addresses unanticipated team member/subteam failures:

- Monitoring and replanning capabilities
  - Explicit constraints of individual/subteam roles and team goal

    AND: All roles must be fulfilled

    OR: At least one role must be fulfilled

    Role-dependency →: Role R1 is dependent on R2
  - Constraints may be combined, e.g. (A OR B) AND (B→C)

Scouting failure example:

- *Wait-for-battle-position-scouted* is the team plan
  - *AND-combination*: Scout and Non-scout roles in team plan
  - If scout crashes, the scout role is not fulfilled
  - AND-combination implies that the team plan fails

STEAM: Monitor and Repair

- *Joint commitment* to replan by reorganization, if critical failures
  - Determine candidates for roles via capability matching
  - Candidates for roles ensure no conflicting critical commitments
  - Individual/subteam may volunteer
  - Highest capability agent wins

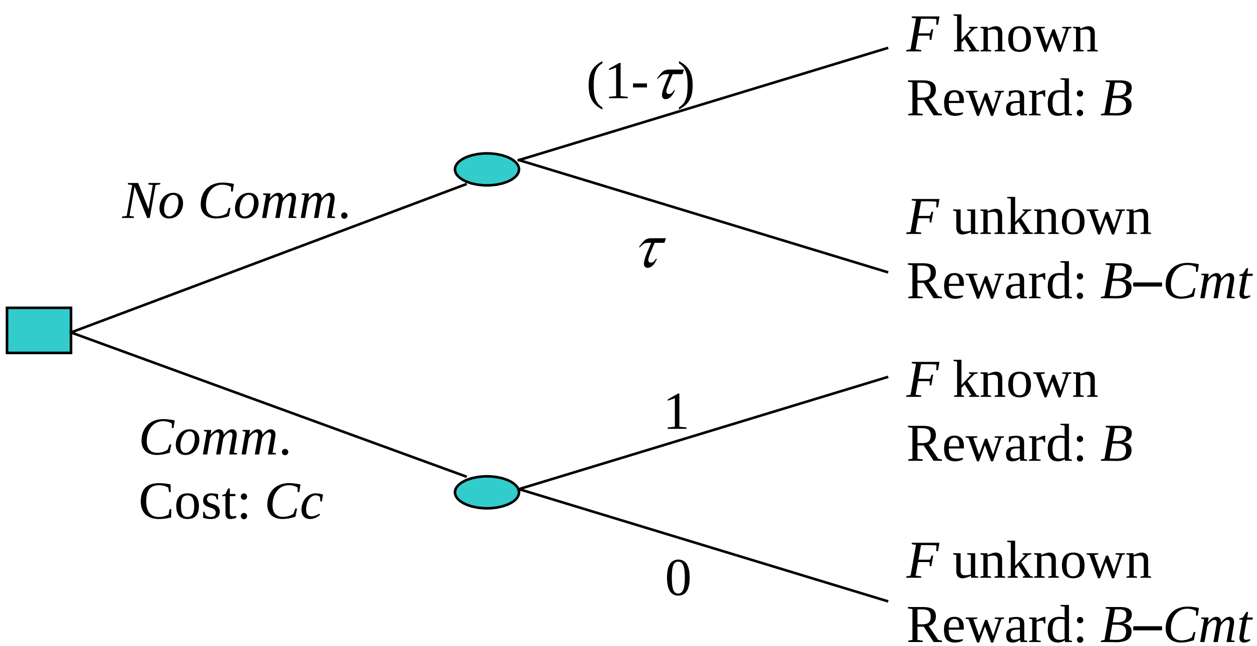
<u>Scouting failure example</u> (continued):

- *Wait-for-battle-position-scouted* is the failed team plan
  - Locate other pilots capable of scouting
  - New candidate scout ensures no conflicting commitments
  - Candidate scout(s) volunteer
  - Best capability scout wins

STEAM: Communication Selectivity

Addresses communication costs and risks:

- Incorporate *decision-theoretic* communication selectivity
- *Tradeoff* communication *vs.* team incoherence costs



EU(No Comm.):
$B - \tau * Cmt$

EU(Comm.):
$B*1+0*(B-Cmt)-Cc$
$=B-Cc$

Communicate if $EU(C)>EU(NC)$: $\tau * Cmt > Cc$

- More complex reasoning being explored: MDP-based models

STEAM: Algorithm Overview

Execute team-orient program (α: team plan, Θ: team(α), π: parameter)

- If *EU*(*C*)>*EU*(*NC*), execute establish-joint-commitment protocol
  - Install team plan α as a joint commitment with team Θ
  - While team-state(Θ) does not satisfy *AUI* conditions of α
    - If F$\in$ private state, satisfies *AUI* conditions of α

      If *EU*(*C*)>*EU*(*NC*), Communicate (Terminate α, Reason: *F*)
    - If receive message(Terminate α, Reason: *F*), update team-state(Θ)
    - If child plan β_i applicable and team(β_i)=Ω and self$\in\Omega$

      In parallel execute team-oriented-program (β_i, Ω, ...)

      Instantiate role constraints

      失败. Recover
    - If teammate failure, evaluate role constraints
- If status(α)=*unachievable*, team-oriented program (*Repair*, Θ, α)

STEAM: Implementations

- Implemented in the *Soar* integrated agent architecture

Example of rules:

Create-communicative-goal-on-achieved

IF

Agent A1's private state contains a fact F AND

Fact F matches an achievement condition of team plan AND

Fact F is not in team state currently

THEN

Create possible *communicative goal* CG to communicate fact F to the team

STEAM: Evaluation and Applications

Applications:

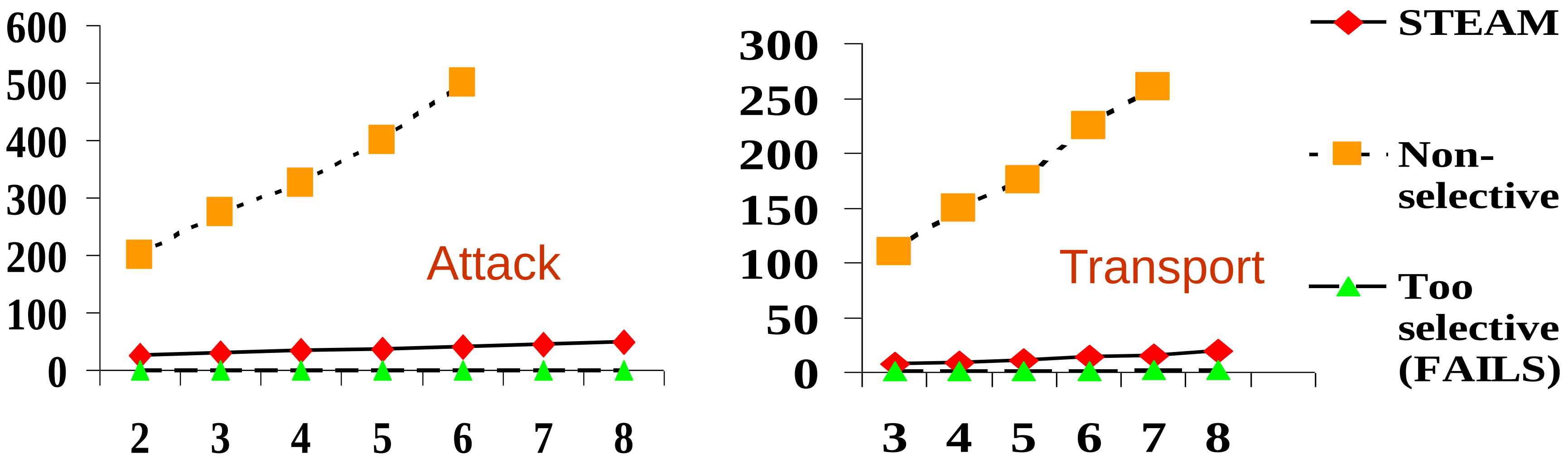
- Pilot teams for synthetic *Attack* agents: *1576 rules*
- Player teams for *RoboCup* Soccer agents: ~ *550 rules*
- Pilot teams for synthetic *Transport* agents: *1333 rules*
- Heterogeneous teams in *TEAMCORE*: ~1000 rules

- STEAM reuse in (% of rules) in complex domains:

Attack	RoboCup	Transport	Teamcore
First use	45%	100%	100%+

STEAM: Evaluation on Communication

- Communication selectivity:



Influence of Teamwork Theories on STEAM

- *Joint commitments* as basic building blocks of teamwork
  - Commitments guide coordination and communication
- Specify entire hierarchy of intentions as in *SharedPlans*
  - Joint intentions for high-level team goal and joint actions for goal attainment
- Knowledge of other team members as in *SharedPlans*
  - Domain-specific coordination relationships explicit to monitor
  - Role-performance ability of individual/subteam monitored
- Joint intentions to *replan* on *unachievability* (*Partial SharedPlans*)
  - Individuals may volunteer or request to take up tasks

GRATE* (Jennings 95)

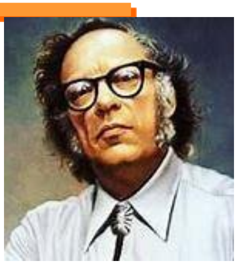
- Cooperative problem solving in electricity transportation domain
- One of the *first* to operationalize joint intentions
- Use *joint responsibility*, based on joint intentions
- Joint-goals as well as joint recipe

- Differences with *STEAM*
  - Two levels of plan hierarchy
  - Three agents (no organization hierarchy)
  - No recovery from agent *failure* (e.g., if one agent cannot fulfill its role)
  - No *selective* communication

Summary: Teamwork

- Teamwork is increasingly *critical* in a vast range of applications
  - Training, education, entertainment, manufacturing, design...
- Fundamental *understanding* of teamwork is important
  - Teamwork theories
  - Not all aspects of teamwork covered
- Practical *teamwork models* based on theories
  - Reuse teamwork capabilities
  - Flexibility in teamwork
  - Team-oriented programming

Asimov's Three Laws of Robotics



- A robot may not injure a human being, or through inaction, allow a human being to come to harm
- A robot must obey the orders given it by human beings except where such orders would conflict with the first law
- A robot must protect its own existence as long as such protection does not conflict with the first and second laws

Isaac Asimov, first published in 1942

Outline

- Understanding teamwork
- Teamwork theories
- Practical teamwork model (STEAM)
- Agent communication
  - Speech act theory（言语行为理论）
  - Formal semantics for speech acts
  - Agent communication language (ACL)

Outline

- Understanding teamwork
- Teamwork theories
- Practical teamwork model (STEAM)
- Agent communication
  - Speech act theory（言语行为理论）
  - Formal semantics for speech acts
  - Agent communication language (ACL)

Speech Act Theory

- Most treatments of communication in (multi-) agent systems borrow their inspiration from *speech act theory*
- Speech act theories are 语用的 *pragmatic* theories of language, — theories of language use
  - Attempt to account for how language is used by people every day to achieve their goals and intentions
- The origin of speech act theories are traced to Austin's 1962 book: *How to Do Things with Words*
- Austin points out that our 话语 utterances function like *physical actions* that can *change the state of the world*
  - Example: "I hereby declare you husband and wife"

Speech Acts

话语

- *Speech act theory* specifies how utterances are used to achieve intentions and goals of the speaker
- Different aspects of *speech acts*:
  - *Locutionary act*（言内行为）

    Speech acts are linked together by grammatical conventions so as to say something meaningful
  - *Illocutionary act*（言外行为）

    In uttering meaningful language, speech acts are used to do things, e.g. provide information, offer advice or make a promise
  - *Perlocutionary act*（言后行为）

    Speech acts have an effect on those who hear the meaningful utterances, e.g. frighten a child by telling a ghost story at night

言语行为理论

- 日常语言哲学中最重要的流派之一、语用学研究中最重要的方法之一，对言语分析影响深远的理论
- Austin (1962) 论著"***How to Do Things with Words*** "《如何以言行事》
  - 人们通过言语完成某种行为（即**言语行为**）
  - 言语行为分为三类：

    言内行为（Locutionary act）：以言表意

    "I hereby declare you husband and wife."

    言外行为（即言外之意，Illocutionary act）：以言行事

    言后行为（Perlocutionary act）：以言取效
  - 其中，**言外行为**是核心，通过断定词、阐释词、执行词、行为词、承诺词等行事动词引入

Classification of Illocutionary Acts

- Searle (1969) identified five *main types* of speech acts:
  - *Assertives*（断言/阐述类）

    such as *informing*, e.g. “It is raining outside”
  - *Directives*（指令类）

    attempt to *get the hearer* to do sth e.g. “please make the tea”
  - *Commissives*（承诺类）

    commit the speaker to doing sth, e.g. “I promise to… ”
  - *Expressives*（表达类）

    express psychological state, e.g. emotion or attitude
  - *Declaratives*（宣告类）

    such as *declaring* war, *designating* officer

言外行为的分类

- Searle (1969) 系统完善了言外行为的分类：

类别	适从方向	典型描述	基本特征
断言/阐述类（Assertives）	话语→现实	表达说话人相信正确的话	陈述句
指令类（Directives）	现实→话语	表达说话人的要求、愿望 如命令、请求、建议等	祈使句、疑问句
承诺类（Commissives）	现实→话语	说话人承担做某事的义务 如承诺、保证、威胁、警告	第一人称
表达类（Expressives）	无适从方向	表达说话人的情感、态度 如祝贺、感谢、抱怨、道歉等	表达说话人的心理状态
宣告类（Declaratives）	双向	通过话语去改变一种情形 如宣战、命名、放弃权利等	说话同时即是动作执行

Speech Act Representation

Speech act representation generally has two parts:

- *Performative verb*（行事动词）
  e.g., request, inform, promise, …
- *Content*
  e.g., "the door is closed"

<u>Examples</u>:

- Speech act: "*Please close the door*."
  - Performative: *Request*, Content: *Close the door* (action)
- Speech act: "*The door is closed*."
  - Performative: *Inform*, Content: *The door is closed* (proposition)
- Speech act: "*Is the door closed?*"
  - Performative: *Inquire/Question*, Content: *The door is closed*

Outline

- Understanding teamwork
- Teamwork theories
- Practical teamwork model (STEAM)
- Agent communication
  - Speech act theory（言语行为理论）
  - Formal semantics for speech acts
  - Agent communication language (ACL)

Plan-Based Semantics

- Agent-based systems require agent communication having a *clearly defined semantics*
- Cohen & Perrault (1979) defined semantics of speech acts using the *precondition-delete-add* list formalism in AI *planning*
- Note that a *speaker* cannot (generally) *force* a *hearer* to accept some desired mental state
- In other words, there is a discrepancy between the *illocutionary act* and the *perlocutionary act*

Plan-Based Semantics

- The semantics for "*request*":

Request (s, h, ϕ)

Preconditions:

- s believe h can do ϕ (i.e. you don't ask someone to do something unless you think they can do it)
- s believe h believe h can do ϕ (i.e. you don't ask someone unless *they* believe they can do it)
- s believe s want ϕ (i.e. you don't ask someone unless you want it)

Effects:

- h believe s want ϕ (i.e. the effect is to make them aware of your desire)

Semantics for Communicative Acts

- Agents build, maintain and disband teams through their performance of *communicative acts* (e.g. *attempt*, *refuse*, *confirm* …)
- Smith & Cohen (1996) defined *semantics* of basic communicative acts for *agent team communication*
- *Communicative acts* serve as the means that agents use to *communicate their mental states* and *form teams with other agents* to achieve their *goals*

➢ Will not cover in class

Definition 4 *Request*

(REQ x y e a p) $\stackrel{\text{def}}{=}$ (ATT x e ϕ ψ)
where ϕ is:
$\neg$(DONE y a) $\wedge$
(PWAG y x (DONE y a) [PWAG x y (DONE y a) p])
and ψ is :
(BMB y x (PWAG x y [(DONE y a) $\wedge$
(PWAG y x (DONE y a)
(PWAG x y (DONE y a) p))]))

Outline

- Understanding teamwork
- Teamwork theories
- Practical teamwork model (STEAM)
- Agent communication
  - Speech act theory（言语行为理论）
  - Formal semantics for speech acts
  - Agent communication language (ACL)

Agent Communication Language

- We now introduce *agent communication languages* (ACLs) — the standard formats for exchanging message
- The best known ACLs are KQML and FIPA, developed by *DARPA knowledge sharing initiative* and *Foundation for Intelligent Physical Agents*, respectively
- Have similar basic structure, for instance, *KQML* consists of *two* parts:
  - Knowledge Query and Manipulation Language (KQML，知识查询与操纵语言)
  - Knowledge Interchange Format (KIF，知识交换格式)

KQML and KIF

- KQML is an "*outer*" language that defines various acceptable *performatives verbs*

  Example performatives:

  - *ask-if* (i.e. is it true that …)
  - *perform* (i.e. please perform the following action …)
  - *tell* (i.e. it is true that …)
  - *reply* (i.e. the answer is …)

- KIF is a language for expressing *message content*, include:

  - *Properties* of things in a domain (e.g. Dan is a professor)
  - *Relationships* between things in a domain (e.g. Bo is Jon's wife)
  - *General properties* of a domain (e.g. Everyone has a mother)

KIF（知识交换格式）Examples

- The temperature of m1 is 83 Celsius（对象之间的关系）

  (= (temperature m1) (scalar 83 Celsius))

- An object is a bachelor if the object is a man and is not married（定义新概念）

  (defrelation bachelor (?x) := (and (man ?x) (not (married ?x))))

- Any individual with the property of being a person also has the property of being a mammal（概念之间的关系）

  (defrelation person (?x) :=> (mammal ?x))

KQML/KIF Dialogue

- Example KQML/KIF dialogue:

  A to B: (ask-if (> (size chip1) (size chip2)))

  B to A: (reply true)

  B to A: (inform (= (size chip1) 20))

  B to A: (inform (= (size chip2) 18))

  … …

KQML/KIF Critiques

- In order to be able to communicate, agents must have agreed on a *common set of terms*
- A formal specification of a set of terms is known as an *ontology*
- The knowledge sharing effort has associated with it a large effort at defining *common ontologies* — software tools like *ontolingua* for this purpose
- KQML/KIF *critiques*:
  - Large performative set, lack strict constraints, inter-operation issue, missing commissive（承诺类）performatives, etc

FIPA

- Since 1995, the Foundation for Intelligent Physical Agents (FIPA) started work on a program of agent standards — the centerpiece is an ACL
- Basic structure is quite similar to KQML:
  - *Performative verbs*

    20 performatives in FIPA
  - *Message content*

```
(inform
        :sender      agent1
        :receiver    agent5
        :content     (price good20 150)
        :language    s1
        :ontology    hpl-auction
)
```

FIPA Performatives

行事动词	传递信息	请求信息	协商	执行动作	错误处理
accept-proposal			x		
agree				x	
cancel		x		x	
cfp			x		
confirm	x				
disconfirm	x				
failure					x
inform	x				
inform-if	x				
inform-ref	x				
not-understood					x
propose			x		
query-if		x			
query-ref		x			
refuse				x	
reject-proposal			x		
request				x	
request-when				x	
request-whenever				x	
subscribe		x			

“Inform” and “Request”

- *Inform* and *Request* are the two basic performatives in FIPA. All others are *macro* definitions, defined in terms of these

The meaning of inform and request is defined in two parts:

- Precondition

  What must be true in order for the speech act to succeed

- “Rational effect”

  What the sender of the message hopes to bring about

“Inform” Performative

For the “*inform*” performative:

- The content is a *statement*
- **Precondition** is that *sender*
  - holds that the content is *true*
  - *intends* that the *receiver believe* the content
  - does *not already believe* that the *receiver knows* whether content is true or not

“Request” Performative

For the “*request*” performative:

- The content is an *action*
- **Precondition** is that *sender*
  - *intends* action content to be performed
  - *believes receiver* is *capable* of performing this action
  - does *not believe* that *receiver already intends* to perform action

References

- M. Tambe, 1997. *Towards Flexible Teamwork*. Journal of Artificial Intelligence Research, 7:83-124, 1997 (STEAM)
- B. J. Grosz and S. Kraus. *Collaborative Plans for Complex Group Action*. Artificial Intelligence, 86(2):269-357, 1996 (SharedPlans)
- P. R. Cohen and H. J. Levesque. *Intention Is Choice with Commitment*. Artificial Intelligence, 42(2-3):213-261, 1990 (Also see next paper)
- P. R. Cohen and H. J. Levesque. *Teamwork*. Nous, 25(4):487-512, 1991. Special Issue on Cognitive Science and Artificial Intelligence
- J. Searle, 1990. *Intentions and Actions*. Intentions in Communication, Chapter 19. MIT Press, 1990 (We-Intentions)
- N. R. Jennings. *Controlling Cooperative Problem Solving in Industrial Multi-Agent Systems Using Joint Intentions*. Artificial Intelligence, 75(2):195-240, 1995 (GRATE*)

References

- J. Austin. *How to Do Things with Words*. Harvard University Press, 1962
- J. R. Searle. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, 1969
- P. R. Cohen and C. R. Perrault. *Elements of a Plan-Based Theory of Speech Acts*. Cognitive Science, 3: 177–212, 1979
- I. A. Smith and P. R. Cohen. *Toward a Semantics for an Agent Communications Language based on Speech-Acts*. Proceedings of AAAI, 1996
- T. Finin, J. Weber and G. Wiederhold. *Specification of the KQML Agent-Communication Language*. DARPA Knowledge Sharing Initiative External Interfaces Working Group, 1994
- *FIPA Specification Part 2 – Agent Communication Language*. Foundation for Intelligent Physical Agents, 1998

End.