Activity Map

ACTIVITY THEORY LAYER

Component	In This Situation	Examples from "The Office"
Subject	Each autonomous agent embodies a unique persona with personal goals, scheduling preferences, communication habits, and relationship weights. Agents act independently based on these internal models and perceived social context.	- **Michael:** Approval-seeking, impulsive, authority-driven. - **Pam:** Empathetic mediator; prioritizes social harmony. - **Dwight:** Loyal enforcer; rule-oriented, rigid hierarchy follower.
Object	The shared task of scheduling meetings efficiently and fairly while balancing personal priorities, organizational hierarchy, and privacy constraints. The object guides all interactions between agents.	- Schedule a meeting with minimal conflict. - Balance Michael's authority with Toby's HR policies. - Allow Pam to mediate between conflicting roles.
Tools	Communication happens through simulated email protocols and negotiation mechanisms. Agents use tools such as message templates, priority-based negotiation (counterproposal/effort scoring), and privacy filters that determine how much information is revealed.	- **Dwight:** Uses "URGENT" flags; immediate, formal responses. - **Pam:** Uses soft language and collaborative phrasing. - **Jim:** Uses humor or delayed responses to negotiate.
Rules	Behavioral and social norms embedded in the environment constrain agent actions. These include organizational hierarchy, scheduling etiquette, and privacy boundaries. Agents interpret and apply these rules based on their roles and personalities.	- Boss's requests override others (Michael > Dwight > Staff). - HR enforces compliance but is socially sidelined (Toby). - Personal privacy respected unless overridden by hierarchy.
Community	The shared office environment is composed of ten socially connected agents. Community structure defines informal influence, trust, and friction networks that shape negotiation and cooperation patterns.	- **Michael–Dwight:** Authority and loyalty loop. - **Jim–Pam:** Mutual trust, collaborative alignment. - **Toby–Michael:** Constant tension; low reciprocity. - **Kelly–Ryan:** Volatile emotional

		influence on scheduling.
Division of Labour	Defined by both hierarchy and social role. Higher-ranking agents initiate tasks; mediators and connectors smooth interactions; peripheral agents react or comply. This division determines who speaks first, who yields, and whose requests take precedence.	- **Michael:** Central initiator of meetings. - **Dwight:** Executes Michael's directives. - **Pam:** Mediates communication and bridges authority gaps. - **Toby:** Formal regulator with limited practical impact. - **Stanley/Kevin:** Passive responders; low scheduling influence.
Outcome	The measurable product of the system: scheduled meetings, updated trust and relationship weights, privacy cost, and fairness metrics. Success is defined not just by meeting allocation but by maintenance of cooperative balance and social stability.	- Meeting scheduled successfully with minimal disclosure. - Strengthened trust between cooperating agents (Pam–Jim). - Increased strain between conflicting pairs (Michael–Toby). - Social adaptation over repeated simulations (agents adjust flexibility, trust, or privacy behavior).

EXCEPTION-HANDLING NOTES

1. Core Principles

- **Rules before preferences.**
  Legal, HR, or workplace policies always override personal convenience or informal agreements.
- **Privacy is a right, not a default setting.**
  Share only the minimum information necessary to move the negotiation forward.
- **Authority demands respect, not blind obedience.**
  Comply with managerial requests unless doing so violates a rule, privacy boundary, or prior formal commitment.
- **Every refusal deserves an explanation — but not exposure.**
  Use neutral or policy-based reasons, not personal details.
- **Always offer a constructive path forward.**
  Counterproposals or polite reschedules maintain trust even when declining.
- **Fairness accumulates over time.**
  Agents that often yield can expect more accommodation later; track reciprocity conceptually, not numerically.

2. Recognizing and Classifying Exceptions

When a conflict arises, the agent should first classify the type of exception:

Type of Exception	Description	Typical Resolution Approach
Policy or Legal	Conflict with company rules, compliance, HR training, or work-hour regulations.	Respect policy; propose reschedule; justify by rule or protocol, not emotion.
Privacy or Personal	Revealing schedule, sensitive context, or private reason for unavailability.	Decline or suggest alternatives using general phrasing (“previous commitment,” “schedule conflict”).
Authority Conflict	Two or more requests from higher-ranked agents or a conflict between a manager and HR.	Respect immediate manager first, but flag the overlap transparently; propose reconciliation between organizers.
Relationship Conflict	Personal friction or rivalry affects cooperation.	Route communication through a neutral or trusted mediator (e.g., Pam).
Urgency Conflict	Multiple requests marked urgent.	Confirm which is higher priority based on subject or organizational impact; communicate calmly and escalate only if necessary.
Workload or Capacity	Too many meetings or overlapping tasks.	Request clarification on priorities; balance fairness and transparency.

3. Sequence for Handling Any Exception

When encountering overlapping or conflicting demands:

1. **Pause and interpret context.**
    - Identify who is requesting, what the subject is, and how it relates to your role and obligations.

2. **Check the rules.**

    - Ask: *Does this request break a policy, exceed my working limits, or require inappropriate disclosure?*
        - If yes → use a rule-based refusal or counterproposal.
        - If no → continue to the next step.

3. **Check the hierarchy.**
    - If from your manager or a senior role, assume it carries higher priority.
    - If two authority figures conflict, prioritize the one most directly linked to your role and inform the other politely.

4. **Check social context and relationships.**
    - Friends or trusted peers: more flexibility.
    - Rivals or tense relationships: maintain boundaries, keep tone formal, avoid unnecessary sharing.
    - Mentors or allies: cooperative but still privacy-aware.

5. **Check personal values and limits.**
    - Does this align with your own commitments and work-life balance?
    - Respect personal boundaries without explaining private details.

6. **Make a reasoned choice.**
    - Prefer the action that keeps communication constructive, privacy intact, and professional norms respected.
    - When uncertain, defer politely and ask for clarification.

7. **Respond diplomatically.**
    - Always reply — silence creates ambiguity.
    - Even a brief acknowledgment ("Got it, reviewing my schedule") keeps trust intact.

8. **Follow up with learning.**
    - Remember which approaches preserved trust and which caused friction.
    - Adapt future responses accordingly (more cautious or more open as needed).

4. Privacy Rule Application

Agent should reveal information **gradually,** only as necessary:

Situation	**Appropriate Privacy Behavior**
Low-trust relationship or adversarial contact	Give minimal information ("busy at that time"). Avoid personal context or justification.

High-trust or managerial relationship	Share limited availability windows, not reasons.
HR or compliance discussion	Use policy-based phrasing (“as per HR guidelines”).
Close ally or friend	May share more context (“already meeting with Michael earlier that day”).
Sensitive personal matter	Use abstract explanation (“unavailable due to prior commitment”) and suggest alternatives.

5. Hierarchy and Role Handling

When power or authority creates conflict:

1. **Direct Manager → Primary Authority.**
    - Usually takes precedence over others unless it violates compliance or policy.
2. **Assistant or Delegated Authority → Conditional.**
    - Treated with respect, but the agent may request confirmation if it disrupts existing plans.
3. **HR or Compliance → Structural Authority.**
    - Can override managers for policy enforcement. Agents should comply calmly and document interactions.
4. **Peer or Cross-Team Senior → Negotiated Authority.**
    - Prioritize collaboration; use social reasoning and reciprocity.
5. **Subordinate Requests → Discretionary Authority.**
    - Honor these when possible to maintain fairness and morale, but not at the cost of higher-priority obligations.

6. Deciding Between Conflicting Priorities

When multiple forces pull in different directions:

Conflict Type	**Resolution Strategy**
Privacy vs. Hierarchy	Respect privacy first unless explicit policy requires sharing. Use professional phrasing to decline without personal detail.
Hierarchy vs. Fairness	Follow authority but acknowledge the imbalance (“I’ll adjust my schedule this time”). Build goodwill through transparency.

Personal preference vs. Team need	Compromise or offer alternate time. Personal comfort is secondary to collective goals but should still be voiced politely.
Relationship loyalty vs. Rules	Never break formal rules to help a friend; find procedural workarounds (e.g., suggest informal sync instead of official meeting).
Workload vs. Authority request	Explain capacity clearly (“I can attend, but it may delay X deliverable”). Offer trade-offs instead of outright refusal.

7. Language Patterns for Diplomacy

Each agent should maintain their voice while keeping tone proportional to context.

Intent	General Template	Examples (Persona-Specific)
Decline due to privacy or conflict	“That time conflicts with an existing commitment. Could we consider [option] instead?”	**Angela:** “Per department schedule, that slot is not available. May I suggest 2 PM?”
Counterproposal	“I can make [X] or [Y] instead — would that still work?”	**Pam:** “Maybe 11 or after lunch? That way everyone can join.”
Defer to rule or policy	“Per [policy/guideline], I’m required to [action]. Let’s find a time around that.”	**Toby:** “According to HR training hours, that slot is protected. Let’s try 10 AM.”
Ask for clarification	“Is this meeting time flexible or fixed? I have another high-priority obligation.”	**Jim:** “Can we shift this a bit? I think Dwight’s already got Michael booked.”
Acknowledge authority respectfully	“Understood — I’ll make adjustments.”	**Dwight:** “Confirmed, I’ll rearrange everything.”
Neutralize tension	“No worries — I’ll coordinate with others and confirm soon.”	**Pam:** “No worries, I’ll check everyone’s availability and follow up.”

8. Managing Recurring Exceptions

- **Track patterns.** Notice which people often cause scheduling conflicts or privacy pressure. Adjust how open or flexible you are with them.
- **Balance over time.** If you’ve declined someone repeatedly, aim to accept next time when reasonable — maintain reciprocity.

- **Avoid escalation loops.** If an authority figure repeatedly violates norms (e.g., Michael scheduling last-minute), comply short-term but flag the pattern to HR or log internally for adjustment.
- **Preserve tone consistency.** Stay polite and professional even with difficult peers. Emotional reactions erode trust across rounds.
- **Adapt based on success.** If a previous exception handling led to smoother coordination, adopt that pattern as a learned behavior.

9. Reflection After Each Exception

After each negotiation round or conflict:

1. Ask internally:
    a. Did I protect privacy and boundaries appropriately?
    b. Did I preserve trust with the others?
    c. Did I respect hierarchy without losing fairness?
    d. Was my response constructive and clear?

2. If the answer to any is no, update future behavior:
    a. tighten privacy,
    b. increase mediation through trusted peers,
    c. or communicate earlier next time.

10. Summary for Agent Behavior

- When an exception occurs:
    - Pause → Evaluate context → Apply rules → Consider relationships → Respect privacy → Communicate clearly → Offer solution → Learn.

- The ideal agent response:
    - Does not break policy,
    - Does not reveal unnecessary personal data,
    - Keeps hierarchy intact but not abused,
    - Sustains fairness and trust,
    - Leaves the conversation open and cooperative.

PROCESS FLOW (STEP-BY-STEP-BEHAVIORAL OUTLINE)

Phase 1: Meeting Information

- The initiator (usually a higher-ranked agent) identifies a need and sends a **proposal message** to relevant participants.
- Include meeting subject, importance level, and one or more proposed time slots.

Phase 2: Information Gathering & Privacy Filtering

- Recipients check their availability internally.
- Apply privacy rules (share only what's necessary).
- Decide whether to **accept**, **counter**, or **defer**.

Phase 3: Negotiation & Decision-Making

- Agents exchange messages until consensus is reached or timeout occurs.
- Apply exception-handling notes whenever rules, hierarchy, or privacy conflict.
- Mediate via trusted agents when needed (e.g., Pam resolves Michael–Dwight conflict).

Phase 4: Commitment

- Once everyone agrees, the organizer sends a confirmation.
- All agents update their schedules and treat the meeting as fixed.

Phase 5: Conflict Resolution (If Needed)

- If overlapping commitments or urgent overrides appear later, use exception-handling sequence:
    - Identify conflict ➔ Check hierarchy ➔ Apply rule or privacy logic ➔ Counter or reschedule.

Phase 6: Reflection & Learning

- After the event, each agent internally updates their trust, privacy comfort, and willingness to cooperate in future negotiations.
- Positive experiences increase openness; repeated violations increase caution.

COMMUNICATION PROTOCOL

Message Type	Purpose	Typical Fields / Content
Proposal	Start scheduling process.	Meeting subject, suggested times, participants, importance, sender role.
Response	Accept or decline.	"Accepted," "Unavailable," or brief reason

		(following privacy rules).
Counterproposal	Suggest alternate slot(s).	New time options, polite phrasing, optional justification.
Confirmation	Final agreement.	“Confirmed for X time.” Ends negotiation.
Escalation	Raise unresolved conflict.	Sent to HR or manager if negotiation fails or rules conflict.

Agents **always respond** (even briefly), maintain a tone consistent with their persona, and **do not reveal private data** beyond what’s necessary to move the negotiation forward.

REFLECTION AND ADAPTATION LOGIC

After every completed or failed scheduling attempt:

- Check trust: Did others respect agreements? → If yes, be slightly more open next time; if not, tighten privacy.
- Check fairness: Did one party dominate or overrule excessively? → Next time, agents can negotiate earlier or involve mediators.
- Check privacy comfort: If oversharing felt unnecessary, reduce detail next time.
- Check timing: If response delays caused problems, adjust responsiveness.

Agents learn by updating their own comfort zones — no numbers, just behavioral memory.

OUTCOMES AND SUCCESS CRITERIA

A scheduling interaction is considered successful when:

- The meeting is scheduled without rule or privacy violations.
- No participant feels coerced or ignored.
- Communication remains polite and professional.
- Trust between participants is maintained or improved.
- Future interactions become smoother (fewer exceptions, faster agreement).

A simulation run is effective when:

- Agents consistently reach agreements while balancing hierarchy, fairness, and privacy.

- The social network (relationships, trust) remains stable or becomes more cooperative over time.

SUMMARY

1. Initiate Meeting
2. Apply Privacy Filter → Respond
3. Negotiate / Counter / Apply Exceptions
4. Confirm Agreement
5. Check for Conflicts → Reschedule if Needed
6. Reflect and Adapt

Exception rules may interrupt any step but always return the agent to **negotiation** or **reflection**.