

# **Resource Management: INSTITUTIONS AND INSTITUTIONAL DESIGN**

SOS3508

Erling Berge

## **Mind and Action**

NTNU, Trondheim

Fall 2009

## **The Mind**

- Insights from
  - Introspection
  - Folk psychology
  - Psychology
  - Behavioural economics
- Needed for understanding
  - Praise
  - Blame
  - Punishment

## Intentions, beliefs, desires

- Current models imputing mental states to others are fragile
  - To maximise expected utility we need to know
    - Values attached to each possible outcome
    - Probability assigned to the occurrence of the outcome
    - Answers will depend on how questions are asked
  - Asking about preferences also reveals context dependence
  - Trade-offs among values are highly unstable
  - Statements about beliefs and mental states are often questionable
  - Beliefs in afterlife and martyrdom
  - Experiences of quasi-emotions (no implication for action)
  - Powers of autosuggestion

Fall 2009

© Erling Berge 2009

3

## Most mental states are unstable

- What about unconscious mental states?
  - If causally effective they can be identified by their effects
  - Unconscious prejudice
  - Unconscious emotions show up in body language probably
  - Self-deception is probably not unconscious

Fall 2009

© Erling Berge 2009

4

## Motivations

- From visceral to rational
  - Acting on emotions
    - Visceral fear vs prudential fear
  - Acting on good reasons
    - Cost-benefit considerations
- Interest, Reason, and Passion
  - Interest is the pursuit of personal advantage
  - Reason is about impartiality and the public good
  - Passion is about emotions and the visceral urges, maybe also forms of madness

Fall 2009

© Erling Berge 2009

5

## Freud

- Id, Ego, Superego as subsystems of the mind
  - Id is similar to passions and corresponds to the pleasure principle
  - Ego is the active I and corresponds to the reality principle
  - Superego is the conscience and corresponds to impulse control
- As ego navigates the external world it also has to fight a two-front war against the impulses from the id (pleasure principle) and the punitively severe impulse control exercised by the superego (conscience)

Fall 2009

© Erling Berge 2009

6

## Taking account of consequences

- Consequentialist motivations
  - The consequences following the actions are their sole motivation
  - E.g.: Most economic behaviour
- Non-consequentialist motivations
  - Consequences are irrelevant, the motivation is the action itself
  - E.g.: conscientious objectivist to military service
  - Kantianism: always do what will be best if all did the same
  - Sanctioning of social norms

Fall 2009

© Erling Berge 2009

7

## Four approaches to motivation

← Biology/emotion
Institutions?
← Cognition/ rationality

1	Visceral Pleasure principle		Rational Rationality principle
2	Passions	Reason	Interest
3	Id	Superego	Ego
4		Non- consequentialism	Consequen- tialism

Fall 2009

© Erling Berge 2009

8

## Wants and wishes

- Wanting and wishing
  - Wants require actions by the one with wants
  - Wishes requires some state of the world to obtain no matter how
- States that are essential by-products
  - Some wants cannot be obtained by direct action, or only accidentally so
  - E.g.: desires to forget, to believe, ...
- Push or pull?
  - Rational: comparisons of two situations
  - Emotional: running away from something (fear, shame) or moving towards goals of glory or honour

Fall 2009

© Erling Berge 2009

9

## Conflicts of motives

- Winner takes all (strongest motive wins)
- If compromise is possible the stronger motive has strongest impact
  - What determine strength here?
- Metamotives
  - Cultural hierarchy of values determine strength of motive
  - Passion and interest will often defer to reason, justice, and fairness
  - Sometimes this is after the fact dressing
- Cognitive dissonance theory suggest small differences in motives will be transformed into large
  - This may result in a kind of path dependence in motivations
- Alliances of motives will determine strength of motivation

Fall 2009

© Erling Berge 2009

10

## Self-interest and altruism

- Altruistic motivation is the desire to enhance the welfare of others even at a net welfare loss to oneself
- Altruistic behaviour is no good indicator of altruistic motives
  - Reason or passion may mimic altruism
  - Love of virtue is different from virtue
- Approbativeness (wish to be thought well about by other people) will often have to stand in for altruism
- Shamefulness (wish to not be thought badly about by others) may also do so
- Social norms encouraging may affect behaviour but to obtain approbation behaviour needs to be supererogatory: in excess of the norm

Fall 2009

© Erling Berge 2009

11

## Transmutation and reciprocity

- Transmutation
  - From interest to reason because of self-love (love of esteem and self-esteem)
  - Finding plausible reasons for self-interested behaviour is easy
- Reciprocity
  - From dyadic to generalised reciprocity
  - Applies both positive and negative (punishment)
  - Applies to building reputations
  - Applies in experiments of trust game

Fall 2009

© Erling Berge 2009

12

## Norms

- Moral norms (proactive)
  - Help others in distress, equal sharing, the everyday Kantian norm
  - Unconditional, but affected by what others do
- Social norms
  - Etiquette, revenge, regulating the use of money,
  - Conditional, Triggered by presence and behaviour of other people
- Quasi-moral norms (reactive)
  - Reciprocity, conditional cooperation,
  - Conditional, Triggered by presence and behaviour of other people
- Identifying altruistic motives
  - Action needs to be proactive, not reactive
  - Action is anonymous
- Imputing motives is often tainted by malice
  - Hermeneutic dilemma
  - Conspiracies occurs
  - Public figures do (sometimes) act on good reasons

Fall 2009

© Erling Berge 2009

13

## Myopia and foresight

- Myopia: Scanning the nearby options to choose the one with largest immediate gratification compared to status quo
  - This leads to a local maximum
- Humans can do better by planning ahead
  - Deferred gratification
  - Choosing the fast road rather than the short
  - Time discounting, high rate means future rewards have low value
    - Involves both cognitive and motivational elements

Fall 2009

© Erling Berge 2009

14

## Pure time discounting

- Usually modelled as exponential discounting: present value of one unit  $t$  periods into the future is  $k^t$  where  $k < 1$  is the per period discount factor. This allows consistent planning
- Empirically this is questionable. To allow for everyday changes of mind and many other phenomena hyperbolic discounting is used. For example with a discount factor of  $1/(1+kt)$  for the present value of one unit  $t$  periods into the future
- Choosing the highest present value in a choice between 10 at  $t=5$  and 30 at  $t=10$  will force a switch some time between  $t=3$  and  $t=4$  (figure 6.3)

Fall 2009

© Erling Berge 2009

15

## Pascal's wager

- You should bet on God's existence since even the smallest chance of being right will give eternal bliss thus trumping all earthly pleasure
- Is the present value of eternal bliss finite or infinite?
  - Exponential discounting gives a finite value
    - Hedonistic life style will be preferred
  - Hyperbolic discounting gives an infinite value
    - Even a small probability of being right will suggest a godly life
    - But at any moment in time hyperbolic discounting will indicate that pleasure today is to be preferred
    - "Give me chastity and continence, but no yet." (St. Augustine)

Fall 2009

© Erling Berge 2009

16



## Weakness of will

- A weak-willed (akratic) person
  - Has a reason to do X
  - Has a reason to do Y
  - In the person's own judgement the reason for doing X is weightier than the reason to do Y
  - The person does Y
- What is the causal mechanism?
  - Preference reversal
    - Temporal proximity (hyperbolic discounting)
    - Spatial proximity (cue dependence)
    - Passions
- These ideas may also apply to temporary changes in beliefs

Fall 2009

© Erling Berge 2009

17

## Beliefs

- Belief as certainty: absolute certainty of being right
- Belief as knowing: justified true belief
- Belief as risk assessment: knowing that one may be wrong with some probability
- Belief as resolution to fundamental uncertainty (at most a ranking of probabilities)
- Belief as resolution to fundamental ignorance

There is a strong tendency for intolerance of uncertainty and ignorance flowing from both pride and a universal propensity to impute meaning, pattern, and agency to the world

- Ignorance and motivation leading to certainty begets errors
- Increasing levels of knowledge may make us more confident than data warrants

Fall 2009

© Erling Berge 2009

18

## Assessing probabilities

- Observing frequencies
  - Do not ignore absolute sizes
  - Think about selection biases
  - Heuristics
    - Availability of an event for the mind
    - Representativeness of a small sequence of events
- Subjective evaluations of available information
  - Few possess good judgement,
    - Also few experts
  - Most should learn to distrust intuitions

Fall 2009

© Erling Berge 2009

19

## Magical thinking

- “Cold” unmotivated mistakes (believing in causal links that cannot exist)
- “Hot” motivated mistakes (unduly influenced by desires)
  - Wanting a belief because it is believed to have good consequences. Getting it is difficult
- Beliefs from content
  - Rationalisation (behaviour first then belief)
  - Wishful thinking (belief first then behaviour)
  - Self-deception (belief first then behaviour)

Fall 2009

© Erling Berge 2009

20

## Emotions

- Source of happiness (love) and misery (shame)
- Impact on action
- Impact on belief
- No agreed definition
- Will focus on
  - Cognitive antecedents (triggering by beliefs, often new)
  - Physiological arousal
  - Physiological expression
  - Action tendencies (there is at least a form of incipient behaviour)
  - Intentional objects (emotion is about something)
  - Valence (strength of happiness or misery)
- There may exist qualitative differences to feelings of emotions like shame vs guilt not reducible to their valence

Fall 2009

© Erling Berge 2009

21

## What emotions are there?

- Focus on those important for social life
  - Evaluative emotions
    - Shame
    - Contempt and hatred
    - Guilt
    - Anger
    - Cartesian indignation
    - Pridefulness (own character)
    - Liking (another's character)
    - Pride (own actions)
    - Gratitude
    - Admiration

Fall 2009

© Erling Berge 2009

22

## More emotions

- From a state of affairs
  - Envy (another's deserved good)
  - Aristotelian indignation (another's undeserved good)
  - Sympathy (another's deserved good)
  - Pity (another's undeserved bad)
  - Malice (another's undeserved bad)
  - Gloating (another's deserved bad)
- Joy and grief from what have or will happen to oneself (known with certainty)
- Hope, fear, love, jealousy based on beliefs that have uncertainty as a component
- Disappointments, regrets, [elation, rejoicing] (relief) caused by some kind of counterfactual reasoning

## Happiness

- A gross national happiness product?
  - Economy as a means to achieve happiness
  - Happiness is subjective
  - Will emotional ups and downs go together?
- Emotion generated action tendency as a temporary preference
- Action is then wanted to occur sooner rather than later
  - Impatience: reward sooner rather than later
  - Urgency: action sooner rather than later

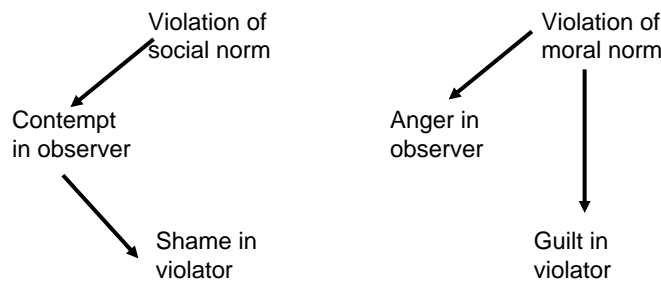
Emotion	Action tendency
Anger or Cartesian indignation	Cause the object of the emotion to suffer
Hatred	Cause the object of the hatred to cease to exist
Contempt	Ostracize; avoid
Shame	"Sink through the floor"; run away; commit suicide
Guilt	Confess; make repairs; hurt oneself
Envy	Destroy the envied object or its possessor
Fear	Flight; fight
Love	Approach and touch the other; help the other; please the other
Pity	Console or alleviate the distress of the other
Gratitude	Help the other

Fall 2009

© Erling Berge 2009

25

## Restoring the moral balance of the universe



- Duration of emotions (short half-life?)
  - Situation dependent, memory dependent
    - But decay may be difficult to anticipate
  - Some are enduring

Fall 2009

© Erling Berge 2009

26

## Emotion and belief

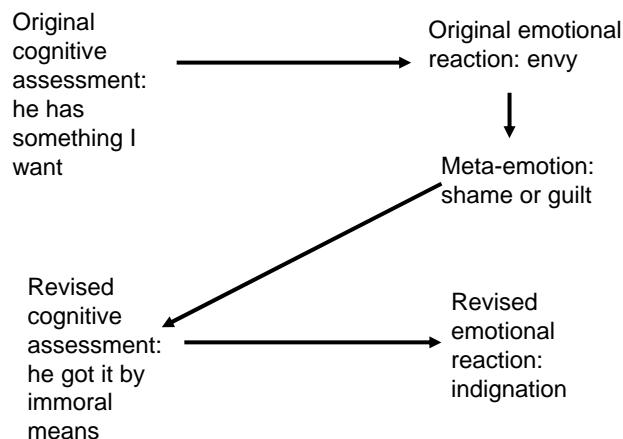
- Direct effects create biased belief
  - Love bias the perception
  - We believe what we fear
- Indirect effects create low-quality beliefs by acting on prior information gathering
- The two often go together reinforcing each other

Fall 2009

© Erling Berge 2009

27

### Transmutation: cognitive rewriting may be necessary



Fall 2009

© Erling Berge 2009

28

## Culture

- Are there universal emotions? Yes, probably all.
- Happiness, surprise, fear, sadness, disgust, and anger have facial expressions recognized across cultures
- The way people think about emotions may be culture specific even if emotions themselves are not
- However, the emotional concepts and ways of thinking about them may also affect behavioural manifestations

Fall 2009

© Erling Berge 2009

29

## Action

- Behaviour (including action (intentional), decision, choice)
  - Not all decisions lead to action
- In choices and decisions we have rationality as our ideal
- This leads to rational choice theory, but not necessarily to more rational actions ...

Fall 2009

© Erling Berge 2009

30

## Desires and opportunities

- Doing one's best
  - Desires define “best”
  - The opportunities frames the choices and give the means (or the beliefs about these do)
  - What is the opportunity set?
  - How can we find the best action in this set?
  - Are opportunities more important than desires?

Fall 2009

© Erling Berge 2009

31

## Opportunities

- They are more easily observed
- Differences in opportunities create variation (indifference curves and budget constraints)
- Are there irresistible desires?
- Opportunities are more easily manipulated than their desires (case: suicide)

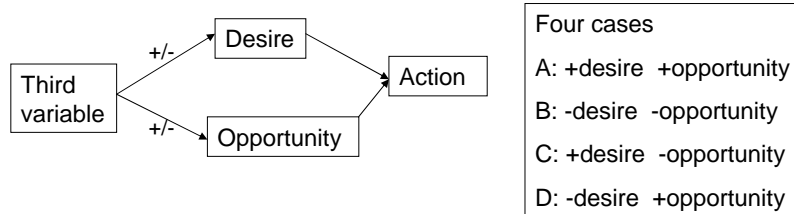
Fall 2009

© Erling Berge 2009

32



## Interactions: simultaneous impacts



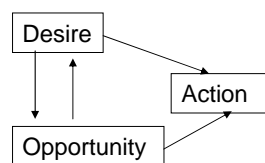
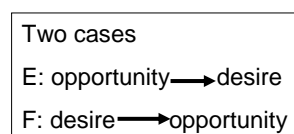
- Explaining action in case B
- Tocqueville:
  - Slavery is unprofitable compared to free labour (negative impact: slavery is not the best choice)
  - Slavery affects the owner's desire to work for his fortune (negative impact)

Fall 2009

© Erling Berge 2009

33

## Interactions opportunity-desire



- E: People end up desiring what the opportunities offer
- F: desiring states that are essential by-products may interfere with the opportunity to get them

Fall 2009

© Erling Berge 2009

34

## Persons and situations

- Can character explain action?
  - Folk psychology assumes stable character and clustering of good traits
    - To some extent this is self-fulfilling
  - To a large extent it is demonstrably false
- The power of the situation
  - Very low consistency of behaviour across situations
  - Consistency of situations leads to disposition of character as explanation: the fundamental attribution error
- No common mechanism can be found in how situations affect behaviour, but both person and situation matter

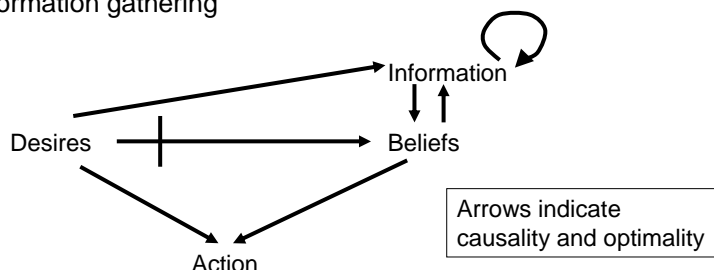
Fall 2009

© Erling Berge 2009

35

## The structure of rational action

- Assuming agents are rational
- Rational action requires
  - The action must be optimal given the beliefs
  - The beliefs must be as well supported as possible given the evidence
  - The evidence must result from an optimal investment in information gathering



Fall 2009

© Erling Berge 2009

36

## Weber was mistaken

- Weber: Departure from rational action is a sufficient indicator to identify irrationality
- But seemingly rational action is not sufficient to certify rationality
- Adaptive reactions (fear and flight) are not rational in the sense used here even if a rational decision making agent would come to the same conclusion

Fall 2009

© Erling Berge 2009

37

## Preferences and ordinal utility

- The desire for the best or optimal outcome
  - Preferences define a rank order of alternative outcomes in terms of “betterness”
- Desires do not have to be “selfish”
  - Rationality does not mean egoism
- Desires do not have to be stable
- Assumptions about preferences (disregarding that the set of options be compact and closed)
  - Preferences have to be transitive
  - Preferences have to be complete
  - To get utilities preferences must be continuous
- Then choosing the best feasible option means maximising utility

Fall 2009

© Erling Berge 2009

38

## Cardinal utility and risk

- Risky choice: action with more than one outcome
  - Choice: option with maximum expected utility (utility times probability)
  - Does not work with ordinal utilities
  - Needs “cardinal utility” (utility measured on an interval or ratio scale)
  - Cardinal utility functions are linear in probability
- Defining cardinal utility
  - A is best option, utility is set to 1
  - B is worst option, utility is set to 0
  - C is some option in between
  - There is a probability  $p(C)$  such that an agent is indifferent between getting C with certainty and a lottery getting A with probability  $p(C)$  and B with probability  $[1 - p(C)]$
  - The cardinal utility  $u(C)$  is defined as equal to  $p(C)$

Fall 2009

© Erling Berge 2009

39

## Risk aversion and decreasing marginal utility

- Risk aversion will conceptually be embedded in the utility function
- Risk aversion may sometimes be confounded with decreasing marginal utility
- Intrinsic utility (impossible to measure so far) is defined by the subjective intensity of enjoyment of some good (intensity of a preference)
- Decreasing marginal utility occurs as the intensity of enjoyment decrease with each additional unit of the good beyond some threshold
- One may also have increasing marginal utility
- Cardinal utility measures the combined effect of risk aversion and intrinsic utility

Fall 2009

© Erling Berge 2009

40

Rational beliefs. Optimal information gathering.

- Rational beliefs are those formed by **processing available evidence by procedures** that in the long run, and on average, are most likely to yield true beliefs
  - Bayesian learning (see p203-204)
- How much time and money should we allocate to acquire new information?
  - Depends on desires (wishful thinking is irrational)
  - Depends on prior beliefs and expected utilities of available options
    - If information costs are above possible gains in utility it is not rational to collect the information
  - Optimal search may depend on the information gathered

Fall 2009

© Erling Berge 2009

41

## Indeterminacy

- Agents may be unable to identify the best element in the feasible set
  - Indifferent between two options
  - Incomplete preferences
  - Unable to determine optimal information gathering
  - Forming beliefs about actions of other agents involved in strategic interactions
    - Reward structure may prevent convergence of beliefs

Fall 2009

© Erling Berge 2009

42

## Rationality is subjective

- Rationality cannot mean good consequences as determined by an external observer. That would be explanation by consequences
- Choices must be seen through the eyes of the agent
  - Choosing a “utility discounting pill” is irrational, the possible actions are already available
- Rational beliefs is different from true beliefs
  - Opportunity costs may discourage investment in information
  - Beliefs about direct costs of information may make investment irrational

Fall 2009

© Erling Berge 2009

43

## Rationality and Behaviour

- Hyper-rationality: optimisation of decisions disregarding the costs of deciding
  - Costs of the means of deciding
  - Costs of side effects of deciding
  - Opportunity costs (value of goods forgone by spending time and resources on deciding)
- Canonical principles of rationality are frequently violated (18 examples discussed)
  - Case 1: Choose acting over non-acting if expected utility of acting is higher
    - Voting can be seen as violating this principle

Fall 2009

© Erling Berge 2009

44

## Alternatives to rational choice theory

- Eleven mechanisms explaining deviations
  - **Loss aversion** (prospect theory)
  - Non-probabilistic weighting of outcomes (prospect theory, utility is non-linear in probability)
  - **Hyperbolic discounting**
  - Heuristics
  - Wishful thinking
  - Inability to project
  - The desire to act for a reason (not the same as acting according to reason)
  - Magical thinking
  - The categorical imperative
  - **Emotions**
  - Social norms

Fall 2009

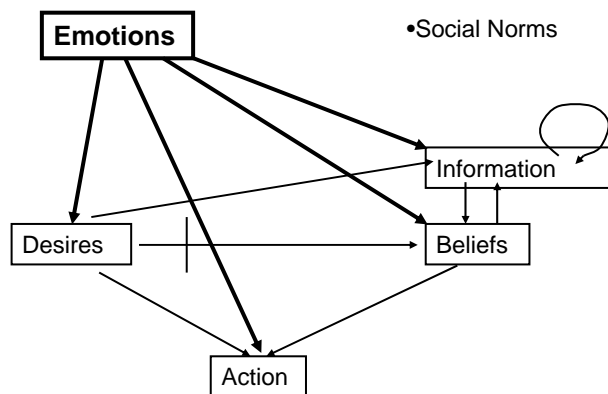
© Erling Berge 2009

45

## Emotions

Emotions may also be involved in

- Magical Thinking and in
- Social Norms



Fall 2009

© Erling Berge 2009

46

## Responding to irrationality

- Second best rationality: taking precautions to avoid irrationalities such as hyperbolic discounting
  - Different from learning
  - Intrapsychic or Extrapsychic (precommitment)
- Agents knowing themselves to be subject to hyperbolic discounting are sophisticated
  - Bundling (or bunching) choices by reframing
  - Acting strategically against a future self
- Extrapsychic devices: precommitment
  - Eliminating a choice of early reward from the feasible set
  - Imposing a penalty on choice of early reward
  - Adding a premium on choice of early reward
  - Imposing a delay between choice and reward
  - Avoiding cues triggering preference reversal
- Precommitment may involve help from other individuals, organisations or public authorities
  - Sometimes this has the form of state paternalism not state assisted self-paternalism
- State constitutions as collective self-paternalism

Fall 2009

© Erling Berge 2009

47

## Implications for textual interpretation

- Interpretations of art (e.g. literature) and explanation are closely related
- A successful work of art can be given a rational choice explanation
  - Explanation by consequences is not allowed
  - Internal (unfolding of plot) vs external explanations (author intentions and dramaturgical know-how): good works can be explained twice over: causally and teleologically
- Rationality requirements for authors
  - Intelligibility of acts and utterances of characters (absolute or relative; global or local)
  - Fullness and parsimony of story (all parts are necessary, non superfluous)
  - The work has to flow “downhill” (plot intelligibility in the minds of readers)

Fall 2009

© Erling Berge 2009

48