

Jon Elster
Tmel společnosti
Studie sociálního řádu

Academia, Praha 2013

Knihla byla vydána s podporou Akademie věd ČR.

KATALOGIZACE V KNIZE – NÁRODNÍ KNIHOVNA ČR

Elster, Jon

Tmel společnosti : studie sociálního řádu / Jon Elster ; [z anglického originálu ... přeložil Rudolf Převrátíl]. – Vyd. 1. – Praha : Academia, 2013. – 357 s. – (XXI. století ; sv. 22)

Název originálu: Cement of society

ISBN 978-80-200-2224-0 (váz.)

316.754 * 316.62 * 331.105.6 * 303.023.5 * 17.021.1

- sociální normy
- sociální chování
- kolektivní vyjednávání
- teorie racionální volby
- jednání
- monografie

3 – Společenské vědy [18]

Published by the Press Syndicate of the University of Cambridge
The Cement of society. A study of social order

© Jon Elster 1989

Translation © Rudolf Převrátíl, 2013

ISBN 978-80-200-2224-0

Obsah

Předmluva 9

Úvod: Dva problémy sociálního řádu 21

1 Kolektivní jednání 29

2 Vyjednávání 65

3 Sociální normy 116

4 Vyjednávání a kolektivní jednání 177

5 Kolektivní jednání a sociální normy 215

6 Vyjednávání a sociální normy 247

Závěr: Tmel společnosti 284

Literatura 328

Jmenný rejstřík 353

Věcný rejstřík 355

Úvod: Dva problémy společenského řádu

Hume napsal ve shrnutí svého *Pojednání o lidské přirozenosti*, že kauzalita je „tmelem vesmíru“. Řád ve fyzikálním světě zajišťuje to, že za událostmi jednoho typu konstantně následují události jiného typu. V této knize se zabývám podmínkami řádu v sociálním světě. Co drží pohromadě společnosti a zabraňuje jim, aby se rozpadly v chaosu a válce? To je velký problém, kterému se co do důležitosti žádný jiný nevyrovná. Netvrdím, že nabízím úplnou odpověď, a ani dílčí odpovědi, které nabízím, nejsou příliš hluboké. Sociální vědy nemohou v současné době aspirovat na víc, než aby byly sociální chemií; induktivními zobecněními, která se drží v těsné blízkosti jevů. Čas pro sociální fyziku ještě nenastal a možná nikdy nepřijde.¹

¹ „Fyzika je skrbtlík. Několik základních idejí platí pro celou přírodu od nepatrnosti atomu až po obrovskou rozlohu galaxie. Navíc tyto základní ideje zachycují v síti svých vzájemných logických vazeb nejrozmanitější faktické informace. Ten, kdo vidí v idejích fyziky půvab a krásu, nemusí nacházet vůbec nic okouzujícího na chemii. Chemikové postrádají jednoduché prediktivní principy, jež jsou pracovní pomůckou fyziků, mají ale úžasnou schopnost držet neustále v hlavě obrovskou sbírku informací. Na druhé straně fyzikové vycházejí ve své práci ze základny tvořené několika málo idejemi, které si je třeba pamatovat.“ (Rigden 1987, s. 36–37).

Budu se zabývat dvěma koncepcemi sociálního řádu: jako souboru stabilních, pravidelných a předvídatelných vzorců chování a jako kooperativního chování. Tomu pak odpovídají dvě koncepce ne-řádu. První ne-řád jako neexistenci předvídatelnosti vyjadřuje *Macbeth* ve své představě života: „Pustý žvást idiota je to, jen hluk a vřava a neznámá nic.“ [Překlad M. Hilského.]² Druhý ne-řád jako nepřítomnost spolupráce je vyjádřen v Hobbesově představě života v přírodním stavu jako „osamělého, ubohého, odporného, hrubého a krátkého“. Ekonomové namísto o předvídatelnosti a spolupráci hovoří o rovnováze a paretovske optimálnosti. Tuto terminologii neakceptuji z důvodů, které se ukáží později. V kapitole 3 dokazují, že sociální normy zajišťují předvídatelnost mimo stav rovnováhy, a v kapitole 5, že spolupráce může vést k výsledkům, jež jsou horší než paretovske optimální.

Dramatickým dokladem ne-řádu jako neschopnosti předpovědi byl propad trhu cenných papírů v říjnu 1987. I když někteří analytici mohli poté právem vyhlášovat „vždyť jsme vám to říkali,“ a dokonce dokazovat, že na základě svých předpovědí skutečně investovali nějaké peníze, podobné Kassandry by bylo pravděpodobně možné najít, i kdyby k havárii došlo o šest nebo čtyřadvacet měsíců dřív. Několik autorů vyslovilo názor,³ že trh cenných papírů je možná *chaotický* režim, a to v technickém smyslu systému, který je charakterizován třemi atributy, jež mohou mít krajně znepokojivé důsledky pro využívání procedur ekonomických předpovědí: a) i když je časová řada generována zcela deterministicky, její chování je statisticky podobné chování systému vystaveného silným náhodným otřesům; b) chaotické časové řady mohou po značně dlouhé časové intervaly vykazovat vzorce chování, jež se zdají být mimořádně uspořádané, načež se objeví bez varování úplně nový vzorec, jen aby stejně nečekaně zmizel; c) výskyt a rozložení takových prudkých přechodů velmi citlivě závisí na parametrech příslušného modelu, přechody se objevují a mizí při změnách

² Tento druh ne-řádu může, ale také nemusí implikovat, že aktéři jsou v nejistotě o tom, co mají dělat. Aktér může mít dominantní strategii, která ho nenechává na pochybách o tom, co dělat, přitom ale nemusí vědět, co chtějí dělat ostatní, nebo o tom bude chovat nesprávné představy. V takovém případě může být skutečným výsledkem *překvapen*, nikdy ale nebude pociťovat *lítost*.

³ Různá interview v práci viz Gleick (1987).

na třetím nebo vyšším desetinném místě, což je naprosto za hranicemi toho, co si může troufat zjišťovat ekonometrie.⁴

Nejsem si ovšem jist, jestli je to správný směr, jakým se mají hledat zdroje nepředvídatelnosti. Nelineární diferenční nebo diferenciální rovnice, které generují chaos, jsou málokdy dobře založeny na mikroúrovni.⁵ Skutečnost, že analytikův model pracuje s chaotickým režimem, nemá velký význam, jestliže není předem teoreticky zdůvodněno to hlavní – důvěra v samotný model. Jestliže navíc z teorie plyne, že model nelze ekonometricky testovat, neexistují ani následně důvody, aby byl brán vážně.

Abychom porozuměli problému nepředvídatelnosti, měli bychom se namísto toho podívat na strukturu sociálního jednání a interakce.⁶ Někdy toho lidé vědí příliš málo o ostatních, aby mohli anticipovat, co udělají, a tudíž předpovědět výsledek. Někdy toho vědí příliš mnoho. Někdy nedokáží využít znalosti, které mají. A někdy jim nemůže pomoci žádné množství znalostí, byť využívaných jakkoli vynalézavě. Dovolte mi ilustrovat tyto čtyři případy, všechny za předpokladu, že lidé jsou racionální. V kapitole 3 se budu zabývat tím, jak může neracionální chování napomoci ke stabilitě a pravidelnosti situací, které by jinak byly beznadějně nestálé.

Lidé mají často značně neúplné informace o racionalitě, preferencích a informovanosti svých protějšků. To nemusí být destabilizující okolnost. Prezident Spojených států neví, jestli by se vůdcové Sovětského svazu chovali racionálně za nějaké krize – jestli by se například zdrželi provedení hrozby, jež by v té době nebylo v jejich zájmu. Nezná preference svých sovětských protějšků (Chtějí doopravdy hegemonii? Nebo jen chtějí být ponecháni na pokoji na svém vlastním dvorku?) A neví také, co se oni sami domnívají o jeho racionalitě, preferencích a informovanosti. Jestliže byla rovnováha teroru přesto stabilní, důvodem pravděpodobně bylo, že obě strany o sobě navzájem předpokládaly tu nejnepříznivější eventualitu a jednaly spíše na základě svých znalostí o objektivních možnostech druhé strany než na nějakých domněnkách o subjektivních duševních stavech. Přes nejistotu a neznalost

⁴ Baumolt a Quandt (1985), s. 3.

⁵ Platí to například o modelech uváděných v pracích Day (1983) a Bhaduri and Harris (1987).

⁶ Viz také Sen (1986).

je zde řád a předvídatelnost. (Jak vysvětlíme později, více znalostí může skutečně způsobit, že bude situace nestabilnější.) Naproti tomu se vztahy mezi Východem a Západem nevyznačovaly uspořádaností v druhém smyslu – ve smyslu kooperativního chování. I když se supervelmoci vyhnuly vzájemnému zničení, dohody o vzájemně výhodném snižování výzbroje jsou nedávného data a jsou co do rozsahu značně omezené.

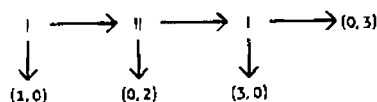
Vzájemná neznalost ovšem často je destabilizující. Uvažujme například o pavučinovém cyklu vyvolaném lidmi jednajícimi za předpokladu, že současné ceny budou platit i v následujícím období. Jestliže současné ceny převyšují rovnovážnou cenu, producenti dají v příštím období na trh objem převyšující rovnovážnou hodnotu, čímž stlačí ceny dolů. Vycházejíce z předpokladu, že nízké ceny potrvají, dají na trh ve třetím období příliš málo, čímž zase ceny zvednou, a tak to půjde dál v cyklu, který může konvergovat k rovnováze, nebo od ní divergovat. Viníkem je zde producentova neznalost spotřebitelů a ostatních producentů. Neví, jak je rozvržena celková poptávka. Neví ani, jak by reagovali jiní producenti, pokud by tento rozvrh znali, protože neví, co se domnívají o sobě navzájem. Za těchto podmínek by mohl stejně tak dobře předpokládat, že ceny zůstanou konstantní bez iluze, že jde o nějaké matematicky podložené očekávání.⁷ Status quo slouží jako těžiště pro utváření přesvědčení. Tak jako hypotéza nejhorší eventuality není předpoklad udržování statu quo racionální ani iracionální. Je to maxima pro rozhodování v nejistotě, která v tomto případě podvrací sama sebe, jestliže ji přijme současně příliš mnoho lidí.

Nadbytek informací může být destabilizující, jestliže překračuje zpracovatelskou kapacitu aktéra nebo organizace, která informace přijímá. V komplexním světě vzniká tento problém často a z praktického hlediska může být docela vážný. Teoreticky se ale o něm mnoho říci nedá. Zajímavější a překvapivější je, že zjevně jednoduché strategické situace mohou být neurčitě, jestliže jsou aktéři plně informováni o sobě navzájem, ale že tyto strategie získají určitost, jestliže informovanost přestane být úplná. Tento případ budu ilustrovat analýzou her v extenzivní formě, jejichž domněle racionální výsledky jsou určovány metodou zpětné indukce.⁸ Jelikož je tato metoda

⁷ Keynes (1936), s. 151.

⁸ Podle práce Stahl (1988) je původcem této metody Zermelo.

jádrem moderního přístupu k iterovaným hrám a k vyjednávání, čímž se podrobně zabývám v následujících kapitolách, budou mít jakékoli pochyby o její platnosti závažné důsledky.



Obr. 1

Namísto toho, abych definoval zpětnou indukci explicitně, budu ji raději ilustrovat na příkladu. Uvažujme o hře mezi dvěma hráči znázorněné na obr. 1.⁹ Je obecně známo, že oba jsou racionální; že oba o sobě navzájem vědí, že jsou racionální, a vědí, že oba vědí, že jsou oba racionální atd. Považka hry je také obecně známá.

Hráči táhnou postupně počínaje hráčem I. Ten může buď táhnout dolů, přičemž získá 1 bod a hráč II nedostane nic, nebo se může rozhodnout pokračovat ve hře ke druhému uzlu a nechat druhý tah na hráči II. Ten může podobně ukončit hru, což mu zajistí 2 body a nic nezůstane pro hráče I, nebo pokračovat ve hře ke třetímu uzlu. V něm má hráč I dvě možnosti, jak ukončit hru. Jedna mu zajišťuje 3 body a druhé straně nic, zatímco druhá má opačný výsledek.

Zdá se být docela jasné, co se stane: I potáhne ve svém prvním kroku dolů a hra tím skončí. Zdůvodnění tohoto závěru se obvykle odvolává na princip zpětné indukce, která postupuje od posledního stadia hry nazpět k prvnímu. Na začátku hry hráč I zvažuje, co by udělal, kdyby byl v posledním uzlu herního stromu. Zřejmě by táhl dolů, nikoliv rovně. Jelikož ví, že hráč II ví, že on sám je racionální, hráč I předpokládá, že hráč II bude v druhém uzlu táhnout dolů, aby získal 2 body a ne 0, což by se stalo, kdyby táhl rovně. Ale toto očekávání nutí hráče I, aby táhl v prvním tahu dolů a získal tak 1 bod nikoliv 0, což by se stalo, kdyby táhl rovně.

Toto zdůvodnění vypadá přesvědčivé. Chová v sobě ale problém: proč by si měl hráč I představovat, že je ve třetím uzlu? Jak by se tam mohl vůbec kdy dostat, jestliže by racionální hráči ukončili hru v prvním uzlu? Věc

⁹ Tento příklad přebírám z práce Bicchieri (1987).

závisí na používání kontrafaktuálních argumentů: za jakých okolností můžeme vyvozovat závěry z premis, o nichž se ví, že jsou nesprávné? Dokazoval jsem na jiném místě,¹⁰ že kontrafaktály lze uplatňovat, jestliže dodatečné premisy použité pro vyvození závěrů z kontrafaktuálního antecedentu jsou slučitelné s tímto antecedentem. Nemůžeme například tvrdit, že pokud by neexistovala železnice, na hospodářském růstu by se mnoho nezměnilo, protože by byl automobil vynalezen dřív, jestliže se toto tvrzení opírá o teorii technických změn, která by nám umožnila předpovědět také vynalezení železnice. Podobně nemůžeme tvrdit, že hráč I bude táhnout dolů, jestliže hra dojde do třetího uzlu, pokud je toto tvrzení založeno na předpokladu racionality hráče II, který je neslučitelný s dosažením tohoto uzlu.

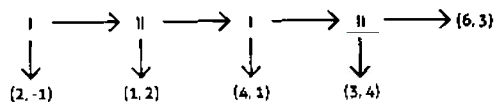
Otázku lze tudíž přeformulovat následujícím způsobem. Existuje soubor předpokladů o hráčích, které budou dostatečně silné na to, abychom mohli vyvodit, co budou dělat hráči v různých uzlech hry, a současně dostatečně slabé, aby byly slučitelné s faktem dosažení těchto uzlů? Dokazoval jsem, že předpoklady racionality a obecné známosti jsou příliš silné na to, aby splnily druhý požadavek. Jiné předpoklady mohou ve zvláštních případech splňovat oba požadavky:¹¹

1. Hráč I by mohl předpokládat, že pozdější uzly byly dosaženy v důsledku chyb (předpoklad „třesoucí se ruky“). Třetí uzel by totiž mohl být dosažen, jestliže by existovala určitá pravděpodobnost, že by některý hráč mohl takřikajíc stisknout nesprávné tlačítko. Jestliže si představíme podobnou hru jako na obr. 1, ale rozšířenou na stovku postupných tahů, byl by takový předpoklad krajně nevhodný, protože by dosažení posledního uzlu vyžadovalo, aby každý hráč udělal padesát navzájem nesouvisejících chyb.

2. Druhým předpokladem proto je, že chyby navzájem souvisejí, například že hráč II je iracionální. Z hlediska hráče I kládoucího si otázku, jak by se dal dosáhnout třetí uzel, je rozumnější předpokládat, že hráč II táhne vždy rovně (předpoklad „automatu“) a že on sám došel do konečného uzlu tak, že využil iracionalitu hráče II.

¹⁰ Elster (1978), kap. 6.

¹¹ Následující úvaha čerpá z práce Binmore (1987a), která shrnuje práce Seltena, Krepsa a Wilsonem a dalších, viz také Rubinstein (1988a) s pronikavými poznámkami.



Obr. 2

3. Hráč I by mohl zvážit složitější možnost, totiž že hráč II je buď iracionální automat, nebo racionální hráč vědomě předstírající, že je iracionální, aby přiměl hráče I k tahu vodorovně a ne dolů. Ve hře znázorněné na obr. 1 se tento předpoklad uplatnit nedá, ale ve hře na obr. 2 může poskytnout hodnověrné vysvětlení toho, že na herním stromu byly dosaženy pozdější uzly.

Aby hráč I vysvětlil, proč by mohl dojít do třetího uzlu, může přisoudit pravděpodobnost p tomu, že hráč II je automat, který vždy táhne vodorovně, a pravděpodobnost $1-p$ tomu, že je racionální aktér předstírající automatické chování v ranějších fázích hry, aby přiměl I k tahu vodorovně, který pak jemu umožní v pozdějších stadiích táhnout dolů.

4. Můžeme zeslabit předpoklad, že racionalita hráčů je obecně známa.¹² Ve hře znázorněné na obr. 1 budeme předpokládat pouze to, že hráč I je přesvědčen, že hráč II je racionální, že hráč II je přesvědčen, že hráč I je racionální a že hráč I je přesvědčen, že hráč II je přesvědčen, že hráč I je racionální. Když si hráč I představí sám sebe ve třetím uzlu, musí si položit otázku, jestli je tento předpoklad slučitelný se souborem uvedených přesvědčení. Konkrétně: mohl hráč II táhnout ve druhém uzlu vodorovně? Aby si hráč I mohl zodpovědět tuto otázku, musí nejdřív vysvětlit, jak si může hráč II smyslu- plně představit, že je v druhém uzlu. V důsledku předpokládané struktury přesvědčení to hráč I nemůže vysvětlit pomocí předpokladu, že hráč II je přesvědčen, že hráč I je iracionální, může ale předpokládat, že hráč II je přesvědčen, že on sám je přesvědčen, že hráč II je iracionální, protože v takovém případě by hráč II mohl usoudit, že hráč I táhne v prvním uzlu vodorovně v naději, že hráč II by mohl udělat totéž v druhém uzlu. Hráč I ale ví, že pokud bude dosažen druhý uzel, hráč II ve skutečnosti bude táhnout dolů, jelikož hráč I je přesvědčen, že hráč II je o něm přesvědčen, že je racionální a tudíž očekává, že hráč I bude ve třetím uzlu táhnout dolů. V souladu s tím hráč I táhne dolů v prvním uzlu. Argumentace zpětnou indukcí funguje, ale

¹² Tento argument uvádí Bicchieri (1987).

jedině proto, že racionalita hráčů *není* obecně známa. Jestliže hráči toho o sobě na začátku vzájemně vědí méně, mohou si vytvořit stabilní přesvědčení o chování svého protějšku.

Předpoklady 1 až 4 se liší tímto: Předpoklad 1 říká, že hráč I se subjektivně domnívá, že existuje určitá pravděpodobnost, že ačkoliv je hráč II racionální, je chybný. Předpoklad 2 říká, že hráč I se subjektivně domnívá, že existuje určitá pravděpodobnost, že hráč II je iracionální automat. Předpoklad 3 říká, že jestliže je hráč II skutečně racionální, bude si vědom hodnocení pravděpodobnosti prováděného hráčem I a pokusí se toho využít ve svůj prospěch, a že hráč I ví, že jestliže je hráč II racionální, bude si toho vědom atd. Předpoklad 4 říká, že hráči jsou racionální a nechybní, ale že jejich vzájemná znalost o racionalitě toho druhého a znalost o této znalosti atd. má horní mez. Každý z těchto předpokladů může ve zvláštních případech a za zvláštních podmínek poskytnout to, co hledáme: teorii dostatečně silnou na to, abychom mohli vyvozovat, co udělají racionální hráči v různých uzlech, a dostatečně slabou, aby byla slučitelná s jejich přítomností v těchto uzlech. Žádný z těchto předpokladů ale podle všeho není tak prostý a obecný, aby mohl potvrdit zpětnou indukcii v obecnějším kontextu. V důsledku toho je tento princip zranitelnější, než se na první pohled zdá.

Mám dojem, že o zpětné indukcii ještě nebylo řečeno poslední slovo. V následujících kapitolách tento princip podržím, byť s jistými pochybami, zčásti proto, že si nejsem jist, jestli chápu tyto věci dostatečně hluboko na to, abych ho mohl zcela zavrhnout, a zčásti proto, že tento princip může být postačující z hlediska chování, jestliže aktéři nejsou dokonale racionální.¹³ Klamná představa obsažená v argumentaci zpětnou indukcí, podobně jako klamné představy v mnoha jiných kontrafaktuálních argumentacích, je tak subtilní, že normální aktéři zabývající se vlastními záležitostmi si jí nemusí povšimnout. Problém čeká na další teoretické a empirické objasnění.

K selhání předpovědí může dojít také proto, že lidé nedokáží správně využít informace, které mají. Jedná se o iracionalitu a nikoliv neurčitost racionality. Tak například v kampani před prezidentskými volbami v USA roku 1972:

¹³ Má zkušenost mi říká, že lidé školeni v matematice uznávají logickou platnost argumentace [zpětnou indukcí], ale odmítají ji akceptovat jako vodítko pro praktické chování (Selten 1978b, s. 133), viz také Rubinstein (1988b).

v předvečer voleb velká skupina reportérů sledujících McGovernovu kampaň mudrlantsky usoudila, že McGovern nemůže prohrát o víc než 10 bodů. Byli to reportéři tiskových agentur, televizních sítí a předních novin a časopisů. Věděli, že všechny významné průzkumy říkají, že McGovern ztrácí 20 bodů a věděli také, že za posledních 24 let se žádný významný průzkum nemýlil o víc než o 3 %. Na své vlastní oči však viděli nadšené desetitisícové davy, které vášnivě aplaudovaly McGovernovi.¹⁴

Trhy cenných papírů a termínové trhy také připisují nadměrný význam aktuálním informacím a nedostatečný význam informacím o minulosti.¹⁵ K selhání předpovědi může vést i opačná chyba – příliš velká snaha o porozumění minulosti. Budeme-li se nějakou dobu rozhlížet po minulosti, bude obvykle možné najít pro jakýkoli soubor minulých událostí jiný soubor událostí, který s prvním vysoce koreluje. Jestliže je požadována pětiprocentní úroveň významnosti, bude v průměru stačit dvacet pokusů. Je ale možné, že tímto způsobem získané korelace budou nepodložené a pro účely predikce k nepotřebě. „Investiční analytikové platí za přehnané připodobňování tím, že nedokáží poskytnout lepší dlouhodobou předpověď, než dávají tržní průměry.“¹⁶

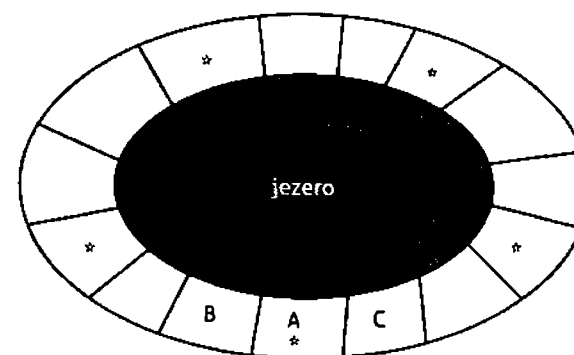
Nakonec konstatujeme, že některé situace jsou inherentně nepředpověditelné. Bez ohledu na to, jak mnoho nebo málo informací mají aktéři a jak důmyslně je využívají, nebudou schopni předpovědět, co budou dělat ostatní. Vycházím z předpokladu, že na to, aby bylo předpovídání racionálními aktéry mezi racionálními aktéry možné, musí být předpovídání výsledek rovnovážným stavem, tj. takovým stavem, v němž žádný aktér nemá pobídku k tomu, aby jednal odlišně. K nemožnosti předpovědi pak může dojít ze tří důvodů: některé situace nejsou v rovnováze, některé mají větší počet rovnovážných stavů a některé mají rovnovážné stavy příliš nestabilní na to, aby mohly sloužit jako základ pro předpověď. První situace není možná ničím víc než kuriozitou,¹⁷ ale zbývající dvě jsou hodně důležité.

¹⁴ Nisbett et al. (1982), s. 116.

¹⁵ Arrow (1982).

¹⁶ Fischhoff (1982), s. 345.

¹⁷ Příkladem je hra, v níž každá osoba napíše číslo a osoba, která napíše největší, získá rozdíl mezi svým číslem a průměrem všech čísel. Ilustrací může být decentralizované mzdové vyjednávání, v němž je všechno dovoleno. Za předpokladu, že firmy mohou



Obr. 3

Neprozkoumatelné mohou být hry s větším počtem rovnovážných stavů, které mají odlišné vítěze a poražené. Představme si jistý počet rolnických rodin s pozemky umístěnými vedle sebe kolem jezera.¹⁸ Rolníci nemají dost orné půdy a každý z nich by chtěl na svém pozemku vykácet stromy, vědí ale, že odlesnění může způsobit erozi. Konkrétně k erozi dojde na jakémkoli pozemku (A na obr. 3) tehdy a jen tehdy, když budou vykáceny stromy na tomto pozemku a dvou sousedních (B a C). Zde jsou tři rovnovážné stavy, při nichž pokaždé zůstanou stromy stát na každém třetím pozemku kolem jezera. Obrázek 3 znázorňuje jeden z nich, při němž zůstávají stát stromy na pozemcích označených hvězdičkou. Neexistuje ale mlčky se rozumějící koordinační mechanismus, jehož působením by se jeden ze tří rovnovážných stavů mohl vynořit jako předpověditelný výsledek.

Jinou situací s větším počtem rovnovážných stavů je hlasování, o němž se podrobněji pojednává v kapitole 5. Jakou hodnotu má hlasování pro jednotlivce, závisí na tom, kolik dalších lidí bude hlasovat. Jestliže hlasují všichni ostatní, jednotlivce k němu nemá žádnou motivaci, protože možnost, že jeho hlas rozhodne, je zanedbatelná. Pokud nebude hlasovat nikdo jiný, má silný motiv,

přesunout jakékoli náklady na spotřebitele, pracovníci ve všech odvětvích nebo firmách mají pobídku, aby požadovali větší zvýšení platů než to, které získali jiní pracovníci. Něco podobného se může stát při hyperinflaci.

¹⁸ Podrobnější rozbor tohoto příkladu uvádí publikace Elster (1989b), kapitoly 9 a 10.

aby tak učinil, protože může sám rozhodnout o výsledku. Rovnováha je dosažena při mezilehlém počtu voličů, z nichž každý dává přednost hlasování před nehlasováním, ale dal by přednost nehlasování před hlasováním, jestliže by hlasovalo o jednu osobu víc. Jestliže je celkem n voličů a rovnováha je dosažena při m voličích, počet rovnovážných stavů je rovný počtu způsobů, kterými se dá vybrat m lidí ze souboru n lidí. Protože m je obvykle malé, volby s velkým počtem oprávněných voličů budou mít velký počet rovnovážných stavů.

Sklon k většímu počtu rovnovážných stavů je značně rozšířenou slabinou modelů racionálního očekávání v ekonomii.¹⁹ Pro jednotlivce neznámá možnost, že může udělat několik předpovědí a každá z nich bude sebenaplňující, žádný problém.²⁰ Může prostě udělat tu předpověď, jejímuž výsledku dává přednost. Před obtížnějším dilematem bude stát skupina, která ví, že pokud budou všichni konat podle jednoho souboru očekávání, stanou se tato očekávání skutečností, a pokud budou jednat podle jiného souboru těchto očekávání, potvrdí se také. Ve zvláštních případech mohou být schopni koordinovat své jednání. Jestliže jeden soubor očekávání vede k výsledku, kterému dávají přednost všichni, bude zvolen. Také mocenské asymetrie mohou stabilizovat situaci a zajistit koordinaci kolem kooperativního rovnovážného stavu. Rolníci kolem jezera mohou dosáhnout koordinaci vyjednáváním, jestliže někteří z nich mají jiné pozemky, které nejsou ohroženy erozí. Protože mohou přežít bez svého pozemku u jezera, mohou věrohodně oznámit, že kácení stromy na své půdě a ostatní se budou muset přizpůsobit jejich jednání.²¹ Pokud ale existuje několik rovnovážných stavů s odlišnými vítězi a prohrávajícími, neexistují obecné mechanismy k zajištění koordinace.²²

¹⁹ Viz např. Begg (1982).

²⁰ Elster (1984), s. 48, 106.

²¹ Tento princip „spravedlnosti podle sv. Matouše“ (tomu, kdo má, bude přidáno) je podrobněji rozebírán v kap. 2.

²² Pokusu utkat se s tímto problémem a vyřešit jej je věnována celá práce (Harsanyi and Selten 1988). Jak ale poznamenává Robert Aumann ve své předmluvě k této knize: „Ačkoliv tato teorie vybírá jedině možný rovnovážný stav, jako teorie nemusí být jedině možná“. Koncepti jejich řešení lze opravdu považovat za představitele jedné z *reflektivních rovnováh* v tom smyslu, jaký jim dává Rawls (1971). Není tudíž jasné, jestli bude mít jejich teorie velkou předpověďací a vysvětlovací schopnost, dokud nebude využita ekonomickými aktéry při jejich rozhodování.

Mnoho her nemá žádný rovnovážný stav při ryzích strategiích. Věrohodným příkladem je investice do výzkumu a vývoje.²³ Jediné rovnovážné stavy, které tyto hry připouštějí, spočívají na smíšených strategiích definovaných jako distribuce pravděpodobnosti nad podmnožinou ryzích strategií. Pravda, nesetkáváme se příliš často s lidmi, kteří při rozhodování v interakcích s nenulovým součtem používají loterie,²⁴ a to z dobrých důvodů. Lze dokázat, že při rovnovážných stavech se smíšenými strategiemi nemůže jednotlivec dopadnout hůř – ačkoliv z definice ani lépe –, jestliže použije jakoukoli jinou pravděpodobnostní kombinaci těch ryzích strategií, jež vstupují do jeho rovnovážné strategie, pokud se ostatní drží svých rovnovážných strategií. Nejneopatrnější závan nejistoty nebo neznalosti pak u něj může způsobit přechod k maximimální strategii: ta ho ochrání, jestliže se ostatní odchýlí od rovnovážného stavu (což udělají, pokud uvažují stejným způsobem), a nemůže ho poškodit, pokud se ostatní neodchýlí.²⁵ Můžeme očekávat randomizaci ve hrách s nulovým součtem, v nichž rovnovážnou strategií je maximimální strategie, ne ale v mnohem důležitějších hrách s nenulovým součtem. Aktér totiž nemůže s jistotou předpovědět, že ostatní přejdou na maximimální strategii, protože pak by mohl využít tuto znalost k volbě strategie, která by optimálně čelila maximimálnímu chování ostatních. V takovém případě by ovšem musel uvážit možnost, že ostatní, kteří se nacházejí ve stejné situaci, budou jednat podobně a tím zruší premisy jeho jednání. Situace je inherentně a ve své podstatě nepředpověditelná.

Uvažujme nyní o neřádu v druhém smyslu, jako o neexistenci spolupráce. To je hlavní téma této knihy. Později uvedeme početné příklady selhání spolupráce a rozlišíme jeho různé příčiny. Zde se omezím jen na několik úvodních poznámek týkajících se klasifikace, přičemž budu rozlišovat mezi

²³ Dasgupta a Stiglitz (1980).

²⁴ Elster v práci Elster (1989a) uvádí v kapitole 2 přehled využívání loterií při přijímání rozhodnutí, viz také Rubinstein (1988a) s některými dalšími kritickými poznámkami k vysvětlovací hodnotě smíšených strategií.

²⁵ A dokonce i tehdy, jestliže je (iracionálně) přesvědčen, že se ostatní budou držet rovnováhy, proč by to měl dělat on sám? Jestliže je skutečně indiferentní vůči všem pravděpodobnostním kombinacím ryzích strategií, proč by měl zvolit právě tu, která uskutečňuje rovnováhu?

pěti hlavními druhy spolupráce, kterými jsou externality, pomoc, konvence, společné podnikání a soukromé uspořádání.

Některé formy spolupráce závisejí na *externalitách* vytvořených individuálním jednáním. Zde má smysl hovořit o individuálních aktech spolupráce: zametání smetí, omezování produkce (jako součásti kartelové dohody), hlasování, placení vlastních daní, darování krve. Lidé nemusí být ochotni spolupracovat v tomto smyslu, pokud neočekávají, že se ostatní budou revanšovat, ale revanšování není nutnou podmínkou účinnosti kooperativního chování. Každý akt spolupráce přináší malý prospěch všem, včetně spolupracujících. Ačkoliv přímé přínosy pro přispívajícího jsou malé na to, aby ho motivovaly k jednání, protože spolupráce s sebou nese také náklady, je pro všechny lepší, jestliže všichni (nebo aspoň někteří) spolupracují, než kdyby nikdo nespolečně pracoval. Za všeobecné spolupráce je každý jednotlivec cílem mnoha malých příspěvků od ostatních, jež v součtu převyšují náklady jeho příspěvku.

Další kategorii nazývám *nápomocné chování*: pomoc sousedovi při sklizení, dodržení slibu, říkání pravdy. Takové jednání nemusí v žádném případě znamenat pro spolupracujícího nějaký přínos. Přínosy se také nešíří na velký počet jiných lidí. Jestliže ale všichni pomáhají ostatním, pokud to od nich vyžaduje jen malé náklady, jsou na tom všichni lépe, než kdyby nikdo nikomu nikdy nepomohl.²⁶ Tak jako v prvním případě je spolupráce ku prospěchu ostatním, ať se revanšují nebo ne. Nemusí k ní ale dojít, pokud se od ostatních neočekává, že laskavost oplátí.²⁷

Třetí je kategorie *konvenčních rovnováh*. Jak už bylo uvedeno, v rovnovážném stavu nikdo nemůže zlepšit svůj výsledek tím, že se odchýlí od rovnováhy. U konvenční rovnováhy si navíc nikdo ani nepřeje, aby se od ní odchýlil kdokoli jiný.²⁸ (Pro konvenční rovnováhy je charakteristické také to, že lidé *striktně* upřednostňují svou vlastní konformitu i konformitu ostatních s konvencí, zatímco běžný pojem rovnováhy vyžaduje jen slabé upřednostňování.) Konvenční rovnováhu vyvolává pravidlo příkazující jezdit po

²⁶ Pomáhání jiným a dodržování slibů (splácení půjček) uvádí Kant mezi příklady kategorického imperativu v *Základech metafyziky mravů*.

²⁷ Viz také Popkin (1979), Elster (1985d), odd. 6.2; M. Taylor, ed. (1988) a samozřejmě Olson (1965).

²⁸ Lewis (1969). Sugden (1986) navrhuje slabší definici.

pravé straně silnice.²⁹ Stejný účinek má pravidlo, kterým se řídí noční život podsvětí v newyorském Bronxu: „Každý, kdo měl přítelkyni, ji bral do společnosti v pátek večer. Nikdo nešel ten večer s manželkou. Manželky chodily do společnosti v sobotu večer. Tak nedocházelo k nehodám, že by narazili na něčí manželku, když byli s přítelkyněmi.“³⁰

Konvence jsou často ekvivalentní: je důležité nějakou mít, ale nezáleží na tom, kterou. Někdy je ovšem některá z možných konvencí lepší v Paretově smyslu než ostatní. I když neexistuje konvence určující, kdo má znovu zavolat, když byl náhodně přerušen telefonický hovor, bylo by zřejmě užitečné nějakou mít. Mohla by určovat, že první bude volající, nebo příjemce hovoru. Z těchto dvou konvencí je zřejmě lepší první, protože jedině volající bude vždy vědět, jaké číslo má volat.

V jiných, významnějších případech mají rozdílné konvence odlišné vítěze a poražené. K tomuto druhu obvykle patří právní systémy. Každý bude mít raději nějaký zákon – v podstatě jakýkoli – než žádný. Ať je zákon jakýkoli, každý ho bude raději dodržovat, pokud správně fungují sankce, a chtějí, aby ho dodržovali všichni ostatní.³¹ Různé skupiny ale mohou mít různé preference, pokud jde o podstatu zákona. Tak například silní i slabí mají společný zájem na zákonech ochraňujících vlastnictví. Slabí dávají přednost režimu, který jim zajišťuje soukromé vlastnictví jejich vlastní pracovní síly a k tomu nějaká vlastnická práva k vnějším objektům, oproti režimu, který jim dává jen zmíněné předchozí právo, dávají ale přednost tomuto minimálnímu režimu před nepřítomností jakéhokoli režimu. Silní dávají přednost minimálnímu režimu, který umožňuje hromadění bohatství, před režimem zajišťujícím slabým širší práva, ale dávají přednost druhému z nich před bezzákonným přirozeným stavem. Z hlediska hromadění bohatství je sociální demokracie účinnější než otroctví, ačkoliv ničím neomezený kapitalismus je ještě lepší. Ovšem silní na rozdíl od slabých by v přirozeném stavu přežili. Mohou

²⁹ Někdy se říká, že *sociální normy* jsou konvenční rovnováhy. V kapitole 3 dokazují, že tato podobnost je povrchní a zavádějící.

³⁰ Pileggi (1987), s. 90.

³¹ Ve skutečnosti si nemyslím, že dodržování zákonů je výlučně funkcí optimálně navržených zákonných sankcí. Zákon je vynucován také sociálními normami. Ty pak mohou být vynucovány sankcemi, které ovšem nejsou zákonné. V kapitole 3 také dokazují, že sociální normy nemohou být plně redukovány na vnější sankce.

tudíž prosadit minimální režim, protože mohou věrohodně vyhrožovat, že se stáhnou do přirozeného stavu.

Další kategorií spolupráce je *společné podnikání*, při němž vytváření přebytku cestou spolupráce vyžaduje fyzickou spolupráci účastníků. V těchto případech nemá jednostranná akce doslova žádný dopad na výsledek. Paradigmatickým příkladem je dělba práce jako v továrně na špendlíky u Adama Smitha. Kolektiv pracovníků vyprodukuje víc společně, než by dokázali odděleně, ale pouze za podmínky, že všichni přispějí svým dílem. Pokud chybí jediný člověk, produktivita všech ostatních klesne na nulu. Naopak člověk navíc nijak neovlivní celkový výkon. Obecněji nebývají v produkci tak striktní komplementarity, ale nadále platí, že kladný marginální produkt daného faktoru vyžaduje nenulovou úroveň každého z ostatních faktorů produkce.

Nádherný příklad dělby práce jako společného podniku podává Garrison Keylor v románu *Lake Wobegon Days*, když popisuje Den vlajky ve svém mytickém městě. Herman, organizátor průvodu, koupil velké množství modrých, červených a bílých čepic a rozdál je lidem. Ti měli projít ulicemi města jako živoucí prapor, zatímco Herman měl stát na střeše Ústřední budovy a fotografovat. Hned po válce lidé spolupracovali velmi ochotně, ale později jim to začalo vrtat hlavou:

Jednou z příčin zlé krve byla skutečnost, že nikdo z nich neměl možnost vidět vlajku, jejíž byl částí – obrázek na papíru byl černobílý. Možnost vidět skutečnou Živoucí vlajku měl jen Herman a pan Hanson a pár hochů příliš malých na to, aby byli zapotřebí tam dole. Lidé chtěli mít možnost vyjít na střechu a sami být svědky té podívané.

„Jak můžete jít nahoru, když máte být dole?“, řekl Herman.

„Půjdete nahoru, abyste se podívali, ale nebudete mít na co. Nestačí vám vědět, že děláte svůj díl?“

Na Den vlajky roku 1949, právě v tom okamžiku, kdy Herman řekl: „To je ono! Teď na moment vydržte!“, se jeden z červených utrl, vyběhl čtyři patra schodů, vyklonil se a dlouze se zadíval. I s dírou, která po něm zůstala, to byl nádherný pohled. Živoucí vlajka docela zaplnila ulici pod ním. Perfektní vlajka! Ta červená je tak zářivá! Nemohl od ní odtrhnout oči: „Maž dolů! Potřebujeme fotku!“ křičel na něj Herman nahoru: „Jak to vypadá?“, křičeli lidé. „Neuvěřitelné! Nedokážu to popsat“, zavolal muž nahoře.

A tak se chtěl podívat každý. „Ne!“, řekl Herman, ale lidé hlasovali a bylo to jednoduše. Jeden po druhém stoupali členové Živoucí vlajky nahoru a obdivovali ji. Byla

úžasná! Vháněla slzy do očí, člověk začal přemýšlet o této velké zemi a místu Lake Wobegon v tom všem. Chtělo se mu stát tam celé odpoledne a vpíjet se do ní. Takže když uplynula první hodina a nahoře bylo pouze čtyřicet lidí z pěti set, ostatních se začal zmocňovat neklid. „Rychleji! Nezdržovat! Už jste to viděli! Běžte dolů a dejte šanci taky ostatním!“ Herman posílal nahoru lidi po čtyřech, pak po deseti, ale po dvou hodinách se ze Živoucí vlajky stala Sedící vlajka a pak se začala rozpadat, protože lidé, kteří se už podívali, mysleli na to, že je čas jít domů na večeři, což rozzuřilo ty, kteří zatím nic neviděli. „Ještě deset minut“, volal Herman, ale z deseti minut bylo dvacet a třicet, lidé se vytráceli a vlajka, která zbyla pro posledního diváka, vypadala jako prostřílená kanonádou.

Roku 1950 převzali Den vlajky Synové Knuta. Herman jim dal krabice s čepicemi. Od té doby Knutové dokázali udělat několik dobrých vlajek, i když po většinu let byla účast slabá. Na dobrou vlajku bylo zapotřebí aspoň čtyři sta lidí. Několik let drželi Knutové pravidlo „žádného divání“, jindy pořádali loterii. Jeden rok experimentovali s velkým zrcadlem, které drželi přes okraj střechy dva muži, ale když se lidé zaklonili a podívali vzhůru, vlajka samozřejmě zmizela.³²

Ve společných podnicích se lidé dohodli na fyzické spolupráci. *Soukromá urovnání* (private orderings) nabízejí vzájemně prospěšné dohody, jež nezávisěji na fyzické spolupráci, ale jen na dobrovolném předání práv za účelem vytvoření přebytku. Jednoduchým příkladem je dvoustranná směna. Tak jako v případě společného podnikání, nemá ani soukromé urovnání místo pro jednostrannou spolupráci, neboť samotný pojem by ztratil smysl. Soukromá urovnání jsou často předehrou ke společnému podniku, předtím, než pracovník smísí svou pracovní sílu s vybavením základními prostředky, je zde smlouva převádějící některá práva k pracovní síle na majitele vybavení.

Řečeno velmi zhruba, teorie kolektivního jednání (kapitola 1) se zabývá prvními dvěma druhy spolupráce, zatímco teorie vyjednávání (kapitola 3) zbývajícími třemi. Teorie kolektivního jednání považuje za hlavní překážku spolupráce problém černého pasažéra. Teorie vyjednávání říká, že hlavním problémem je neschopnost dohodnout se o dělbě výnosů spolupráce. To je ovšem jen velmi přibližné vymezení. Hlavní argument knihy zní, že *neúspěchy kolektivního jednání jsou často způsobeny tím, že selže vyjednávání*. Bylo by často absurdní požadovat, aby všichni přispěli k veřejnému dobru

³² Keillor (1986), s. 123–124.

stejnou měrou. Někteří ho potřebují víc než ostatní, nebo si mohou spíš dovolit přispět. Velikost jejich příspěvku se pak stane předmětem vyjednávání. Někdy nemá příspěví všech význam, nebo je dokonce škodlivé. Pak se bude vyjednávat o tom, kdo přispěje a komu se dovolí svézt se zdarma. V obou případech ke kolektivnímu jednání nedojde, pokud se potenciální přispěvatelé nedohodnou. Tento společný základ teorie kolektivního jednání a teorie vyjednávání je zkoumán v kapitole 4.

Při rozebírání těchto problémů používám dva hlavní koncepční nástroje. Prvním je *teorie racionální volby*, které jsem věnoval značnou pozornost v jiných pracích.³³ Zde se jí nebudu zabývat explicitně kromě toho, že poslouží jako pozadí pro diskusi o druhém nástroji, *teorii sociálních norem*. Dospěl jsem k přesvědčení, že sociální normy poskytují důležitý typ motivace pro jednání, které nelze redukovat na racionalitu nebo jakoukoli jinou formu optimalizujícího mechanismu. Tento názor vyložím a budu obhajovat v kapitole 3. V kapitolách 5 a 6 znovu zvažuji kolektivní jednání a vyjednávání z dodatečného hlediska teorie sociálních norem. Závěrečná kapitola nabízí některé další úvahy a předběžný závěr.

Dokazování je, jak jsem řekl, hlavně induktivní. Značně se opírám o příklady ilustrující a dokreslující abstraktnější tvrzení, která by jinak mohla začít žít tím druhem vlastního života, jenž sužuje sociální vědy jako noční můra. V důsledku této induktivní strategie se může zdát, že výklad je mísy dost nesouvislý. V románu *Melville Goodwin, U.S.A.* od Johna Marquanda někdo o titulní postavě říká, že její příběhy pobíhají po podlaze jako králíci. Nebyl bych překvapen, kdyby si někteří čtenáři mysleli totéž o této knize. Snažil jsem se jí nicméně dát jistou soudržnost tím, že jsem jako hlavního nositele argumentace použil *kolektivní mzdové vyjednávání*. Tento problém je podrobněji vysvětlován střídavě v několika kapitolách, a to 2, 4 a 6. Jak bylo vysvětleno v předmluvě, otázka kolektivního vyjednávání byla vlastně tím prvním, co jsem začal zkoumat. Když se ukázalo, že je příliš obtížná, přešel jsem ke snadnějšímu úkolu vytvoření rámce, v němž by bylo možné tuto otázku zkoumat. V důsledku toho už nemá mzdové vyjednávání a tvorba mezd status explananda, ale skromnější postavení ilustrace.

³³ Elster (1983a, 1984, 1986, 1989a).

Úvod

Problém kolektivního jednání, kterému se také říká problém černého pasažéra nebo problém dobrovolného poskytování veřejných statků, je hluboký a všudypřítomný. Racionální sebezájem může vést jednotlivce k tomu, aby se chovali způsoby, které jsou kolektivně katastrofální. Aby lidé unikli z tohoto dilematu, mohou se vzdát své moci ve prospěch státu, Hobbesova Leviathana. Mohou ovšem dosáhnout spolupráci také decentralizovanými, nevynucovanými prostředky. To je hlavní téma této kapitoly. Decentralizovaná řešení jsou základnější než centralizovaná, protože dodržování centrálních direktiv je samo o sobě problémem kolektivního jednání. V této kapitole upřesňuji strukturu problému kolektivního jednání a rozebírám, jak mohou lidé racionálně chtít spolupracovat a nikoliv svézt se zadarmo s ostatními. Analýza je neúplná, protože do rozhodnutí spolupracovat vstupují silně také neracionální motivy. Ty rozebírám v kapitole 5.

Abych dal podnět k diskusi, začnu s několika konkrétními příklady problému kolektivního jednání. Vytváření odborové organizace nebo cenového kartelu je problémem kolektivního jednání pro potenciální členy: všichni mají prospěch, jestliže všichni vstoupí, ale každý má větší prospěch, jestliže