nutelný. Dokázal, že na *Entscheidungsproblem* existuje odpověď, a tou je „ne". Nevypočitatelné číslo je ve skutečnosti nerozhodnutelným výrokem.

Turingův stroj - fantastický, abstraktní, zcela imaginární počítač - tedy vede k důkazu, který je zrcadlovým obrazem Gödelova důkazu. Turing však zašel dál než Gödel, neboť definoval obecný koncept formálního systému. Každý mechanický postup na vytváření formulí je v podstatě Turingovým strojem. *Jakýkoli* formální systém tedy musí mít nerozhodnutelné výroky. Matematika není rozhodnutelná. Neúplnost je důsledkem nevypočitatelnosti.

Opět zde vidíme, že paradoxy se vynoří, jakmile čísla získají moc kódovat vlastní chování stroje. Je to nutný rekurzivní obrat. Počítané je osudově propletené s počítajícím. Douglas Hofstadter to o mnoho let později vyjádřil větou: „Ta věc závisí na spuštění testu zastavení, jenž ve snaze předpovědět své vlastní chování, když pozoruje sebe sama ve snaze předpovědět své vlastní chování, když pozoruje sebe sama ve snaze předpovědět své vlastní chování, když..."[9] Obdobný hlavolam se nedávno objevil také ve fyzice - nový princip neurčitosti Wernera Heisenberga. Když se o něm dozvěděl Turing, vyjádřil ho jazykem autoreference: „Ve vědě se obvykle předpokládalo, že pokud bychom v nějakém okamžiku věděli o vesmíru vše, pak bychom mohli předpovědět celou jeho budoucnost... Modernější věda však došla k závěru, že když se zabýváme atomy a elektrony, nejsme schopni je přesně poznat - vždyť naše nástroje se skládají z nich samotných."[10]

Mezi Babbageovým úžasným, leč neskladným Analytickým strojem a elegantní, ale neskutečnou abstrakcí s názvem Turingův univerzální stroj uplynulo celé století. Turing se nikdy ani nesnažil o fyzickou realizaci stroje, o němž matematik a logik Herbert Enderton o řadu let později poznamenal: „Můžeme si představit pilného a pečlivého úředníka, který má vždy dostatek papíru na poznámky a neúnavně se řídí pokyny."[11] Turing byl podobně jako Ada Lovelace programátorem, který dopodrobna pozoruje sekvenční logiku vlastní mysli. Sám sebe si představoval jako počítač. Své myšlenkové postupy redukoval až na nejmenší složky, atomy zpracování informací.

Alan Turing a Claude Shannon měli společnou zálibu v kódech. Turing kódoval pokyny do podoby čísel, desítková čísla kódoval jako nuly a jedničky. Shannon vytvářel kódy pro geny a chromozomy i pro relé a přepínače. Oba muži uplatnili svůj důmysl v mapování jedné množiny objektů na druhou - příkladem mohou být logické operace a elektrické obvody nebo algebraické funkce a pokyny pro stroje. Hra symbolů a idea *mapování*, ve smyslu hledání jednoznačné korespondence mezi dvěma množinami, zastávaly přední místo v jejich myšlenkové výzbroji. Jejich kódování nemělo věci zahalovat, ale naopak

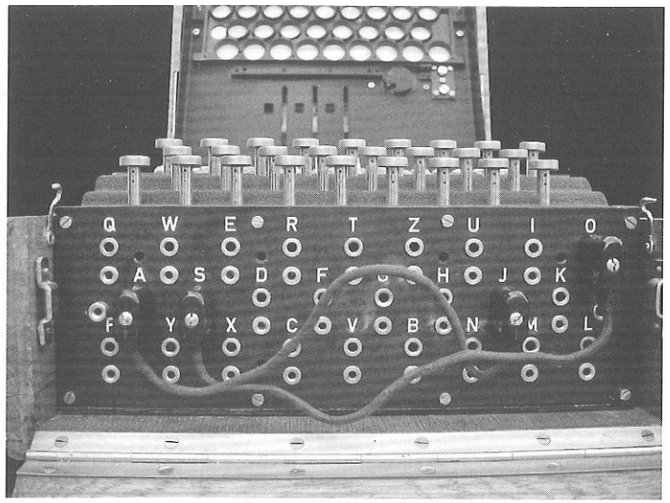
osvětlovat - chtěli ukázat, že jablka a pomeranče jsou koneckonců obdobné, a když ne obdobné, tak aspoň zastupitelné. Válka přivedla oba muže ke kryptografii - úsilí věci naopak co nejvíce zahalit.

Turingova matka se svého syna často ptala, k čemu je ta jeho matematika dobrá. Již v roce 1936 jí odpověděl, že objevil její aplikaci „na mnoho zvláštních a zajímavých kódů".[12] A dodal: „Předpokládám, že bych je mohl za poměrně značnou částku prodat vládě Jejího Veličenstva, ale pochybuji o morálnosti takových věcí." Turingův stroj opravdu mohl *vytvářet* šifry. Ukázalo se však, že vláda Jejího Veličenstva má jiný problém. Úkol číst zprávy, které zachytili z německého kabelového i bezdrátového spojení, přešel v předvečer války na Vládní školu kódů a šifer, jež původně patřila pod Admiralitu. Jejími pracovníky byli zpočátku jazykovědci, úředníci a písaři na stroji, ale žádný matematik. Turing se připojil v roce 1938. Když se škola kódů a šifer odstěhovala z Londýna do bezpečnějšího Bletchley Parku, venkovského sídla v Buckinghamshiru, nastoupil do sekce, ve které byli i šachoví mistři a vítězové soutěží v luštění křížovek. Již bylo jasné, že klasické jazykové vzdělání může kryptografii stěží nějak obohatit.

Německý systém s názvem Enigma používal polyalfabetickou šifru, kterou vytvářel rotorový stroj o velikosti kufříku s připojenou klávesnicí a signálními žárovkami. Tato šifra měla slavného předka ve Vigenèrově šifře, jež byla považována za neprolomitelnou, než ji v roce 1854 prolomil Charles Babbage. Babbageův vhled do matematiky vědecké sekci v Bletchley zpočátku velmi pomohl, stejně jako práce polských kryptografů, kteří měli za sebou roky tvrdé praxe s prolamováním šifer wehrmachtu. Turing pracoval v objektu známém jako „Hut 8", kde se ujal vedení v teoretickém výzkumu a problém vyřešil nejen matematicky, ale i fyzicky.

Znamenalo to zhotovit stroj, který bude umět dešifrovat všechny šifry od Enigmy. Zatímco jeho první stroj byla jen abstrakce s pomyslnou páskou, konkrétní stroj s přezdívkou Bomba zabíral tři krychlové metry, obsahoval tunu drátů a kovu nasáklého olejem a účinně mapoval rotory německého přístroje na elektrické obvody. Triumf vědy v Bletchley - který zůstal tajemstvím po celou dobu války a dalších 30 let po ní - měl ještě větší dopad na výsledek války než Projekt Manhattan, jenž vedl k vývoji skutečné bomby. Ke konci války dešifrovaly Turingovy Bomby tisíce vojenských odposlechů denně - zpracovávaly informace v takovém rozsahu, jaký svět dosud neviděl.

Když se Turing a Shannon setkali u jídla v Bell Labs, sice se o ničem z toho nebavili, ale nepřímo hovořili o Turingově představě, jak všechny tyto *věci* měřit. Turing sledoval, jak analytici zvažují zprávy, které kolovaly po Bletchley - některé nejisté a navzájem si odporující - když se snažili zjistit pravděpodobnost nějaké skutečnosti, ať se jednalo o konkrétní nastavení Enigmy nebo polohu ponorky.



Ukořistěný stroj ENIGMA

Turing měl pocit, že je zde něco, co je třeba matematicky změřit. Nebyla to pravděpodobnost, která se obvykle vyjadřovala jako poměr šancí (například 3 : 2) nebo číslo mezi 0 a 1 (jako 0,6 nebo 60 %). Turing se zajímal o data, jež pravděpodobnost *měnila*: o pravděpodobnostní faktor, něco jako váhu důkazů. Vymyslel jednotku, kterou pojmenoval „ban". Pro práci mu vyhovovala logaritmická stupnice - bany se tak mohly přičítat, a ne násobit. Když byl základ 10, ban byl vahou důkazů nutnou k vytvoření desetkrát pravděpodobnější skutečnosti. Pro jemnější měření Turing používal „decibany" a „centibany".

Shannon uvažoval podobně.

Při práci na starém štábu West Village rozvinul teoretické představy o šifrování, s jejichž pomocí se mohl soustředit na svůj sen, který prozradil Vannevaru Bushovi - „analýzu některých základních aspektů obecných systémů, jež slouží k přenosu zpráv". Po celou válku postupoval souběžně po dvou liniích - nadřízeným ukazoval kryptografické výsledky a zbytek si nechával pro sebe. Utajení bylo heslem dne. Pomocí čisté matematiky Shannon analyzoval i některé ze šifrovacích systémů, na které Turing útočil se skutečnými odposlechy a nejvýkonnějším hardwarem. Například se jednalo o konkrétní otázku bezpečnosti Vigenèrovy šifry v situaci, kdy „nepřítel zná používaný systém".[13] (Němci používali právě tuto šifru a Britové byli tím nepřítelem, který používaný systém zná.) Shannon si všímal nejobecnějších případů. Ty všechny obsahovaly „diskrétní informace", tedy posloupnosti symbolů vybrané z konečné množiny: zejména šlo o písmena abecedy, ale také o celá slova a dokonce i o „kvantovanou řeč" -

hlasové signály, rozdělené do paketů s různou amplitudou. Zašifrovat taková data znamenalo zaměnit jedny symboly za druhé podle určitého matematického postupu, přičemž příjemce zprávy zná klíč, který použije k rekonstrukci zaměněných symbolů. Dokud zůstane klíč utajen, systém zůstane bezpečný i tehdy, když nepřítel zná postup.

Luštitel šifer vidí proud dat, který připomíná blábol, a snaží se najít původní signál. Shannon k tomu poznamenal: „Z hlediska odborníka na dešifrování je utajený systém téměř stejný jako zašuměný komunikační kanál."[14] (Svou zprávu „Matematická teorie kryptografie" dokončil v roce 1945. Okamžitě byla klasifikována jako tajná.) Proud dat má vypadat jako náhodný šum, ale samozřejmě takový není - pokud by byl opravdu náhodný, signál by se ztratil. Šifra musí přeměnit běžný jazyk, který má určitou strukturu, na něco, co ji na první pohled postrádá. Struktura je však překvapivě odolná. K analýze a roztřídění transformací při šifrování musel Shannon pochopit jazyk tak, jak to učenci, například lingvisté, nikdy předtím nezkoušeli. Lingvisté nicméně již začali zaměřovat svoji pozornost na stavbu jazyka - struktury, které lze najít mezi neurčitými, rozevlátými tvary a zvuky. Lingvista Edward Sapir se zmiňoval o „symbolických atomech", jež jsou tvořeny základními fonetickými vzory jazyka. V roce 1921 napsal: „Pouhé zvuky mluvené řeči nejsou základem jazyka, ten spočívá spíše v klasifikaci, ve formálním strukturování ... Jazyk jako stavba je ve své vnitřní podobě šablonou myšlení."[15] *Šablona myšlení* - to znělo slibně. Shannon však potřeboval pro jazyk hmatatelnější a počitatelnější pojmy.

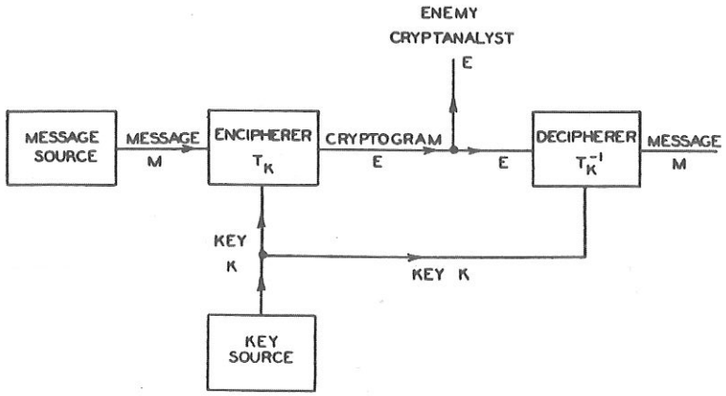
Struktury v jazyce podle Shannona vyplývají z jeho redundance. V běžném jazyce slouží redundance k lepšímu pochopení smyslu, při šifrování je však Achillovou patou. Kde se taková redundance nachází? Například kdekoli se v angličtině objeví písmeno *q* a za ním písmeno *u*, je toto *u* nadbytečné. (Nebo téměř nadbytečné - zcela nadbytečné by bylo, kdyby neexistovalo několik málo přejatých slov jako *qat* a *Qatar*.) Po *q* se *u* zkrátka očekává, není žádným překvapením, nepřináší žádnou informaci. Částečně nadbytečné je také písmeno *h* za písmenem *t*, protože je to nejpravděpodobnější písmeno, které se po *t* může objevit. Shannon tvrdil, že každý jazyk má určitou statistickou strukturu a s ní i určitou redundanci. Navrhl, abychom ji nazvali *D*: „*D* v jistém smyslu měří, nakolik může být text daného jazyka zkrácen bez toho, aby se ztratila jakákoli informace."[16]

Shannon odhadoval, že redundance angličtiny činí kolem 50 %.* Bez počítačů, které by zpracovaly velké objemy textu, si tím nemohl být jistý, ale jeho

* „...aniž by se brala v úvahu statistická struktura na větší vzdálenost než asi osm písmen."

odhad se ukázal jako správný. Typické pasáže se dají zkrátit na polovinu, aniž by se tím ztratila informace. (*If u cn rd ths...*) U nejjednodušších raných substitučních šifer tato redundance znamenala první slabinu. Edgar Allan Poe věděl, že když šifra obsahuje více písmen *z* než jiných písmen, je *z* pravděpodobně náhražkou za písmeno *e*, jelikož *e* je v angličtině nejčastějším písmenem. Jakmile se rozluštilo *q*, rozluštilo se tím i *u*. Luštitel šifer hledal opakované struktury, které mohly odpovídat běžným slovům nebo kombinacím písmen, jako *the, and, -tion*. Ke zdokonalení frekvenční analýzy potřebovali tito luštitelé více informací o četnosti písmen, než dokázali získat Alfred Vail nebo Samuel Morse zkoumáním četnosti liter v zásobníku sazeče. A promyšlenější šifry tuto slabinu stejně překonaly neustálým obměňováním zástupné abecedy - každé písmeno pak mělo mnoho možných náhražek. Zřejmé, snadno rozeznatelné struktury zmizely. Pokud si však zašifrovaný text uchoval nějaké stopy struktur - nějakou formu, pořadí či statistickou pravidelnost - mohl matematik teoreticky najít cestu, jak šifru rozluštit.

Všechny šifrovací systémy měly společnou jednu věc: používání klíče. Bylo jím kódové slovo, fráze, celá kniha nebo něco ještě složitějšího, ale v každém případě šlo o zdroj znaků, který znali odesilatel i příjemce - vedle samotné zprávy to byla další informace, kterou sdíleli. V německém systému Enigma byl klíč obsažen uvnitř technického vybavení a denně se měnil. Odborníci z Bletchley Parku ho museli pokaždé znovu objevit, když zkoušeli texty nově zašifrovaných zpráv. Shannon se mezitím uchýlil k těm nejabstraktnějším, nejobecnějším a nejteoretičtějším východiskům. Tajný systém obsahoval konečný (i když asi značný) počet možných zpráv, konečný počet možných kryptogramů a mezi nimi, kde se jedno měnilo v druhé, také konečný počet klíčů, z nichž každý měl svou pravděpodobnost. Shannonovo schéma vypadalo takto:



Nepřítel i příjemce se snaží dosáhnout stejného cíle, kterým je původní zpráva. Když to Shannon koncipoval tímto způsobem, z hlediska matematiky a pravděpodobnosti, zcela oddělil ideu zprávy od jejích hmotných detailů. Zvuky, tvar vlny, všechny obvyklé starosti technika Bell Labs - nic z toho nyní nehrálo roli. Na zprávu se nahlíželo jako na volbu - jednu možnost vybranou z množiny možností. Když nechal Paul Revere v noci poslat znamení z kostela Old North Church, byly možné jen dvě zprávy (postupují-li Britové po souši nebo po moři). Nyní se téměř nedaly spočítat, přesto však umožňovaly statistickou analýzu.

Shannon se nemusel zabývat luštěním skutečných šifer jako jeho kolegové v Bletchley Parku, nicméně vytvořil rozsáhlou konstrukci z algebraických metod, teorémů a důkazů, jež kryptologům poskytla to, co nikdy předtím neměli - rigorózní metodu vyhodnocení bezpečnosti jakéhokoli šifrovacího systému, čímž položil vědecké základy kryptografie. Kromě jiného prokázal existenci dokonalých šifer - slovem „dokonalých" se myslelo, že ani nekonečně dlouhá ukořistěná zpráva luštiteli nijak nepomůže („i když nepřítel zachytí sebedelší zprávu, nebude na tom o nic lépe").[17] Ovšem na druhé straně dokázal, že kvůli nesmírně přísným požadavkům byly tyto šifry prakticky nepoužitelné. V dokonalé šifře musí mít všechny klíče stejnou pravděpodobnost, musí se tedy jednat o náhodný proud znaků; každý klíč lze použít jen jednou; a vůbec nejhorší je, že každý klíč musí být stejně dlouhý jako celá zpráva.

Ve své tajné zprávě „Matematická teorie kryptografie" Shannon téměř mimochodem použil termín, který nepoužil nikdy dříve: „teorie informace".

Shannon se nejprve musel zbavit „smyslu". (Uvedené „dezinfekční" uvozovky jsou jeho.) Vesele prohlásil: „‚Smysl' zprávy je obvykle bezvýznamný."[18]

Provokoval, aby zcela vyjasnil svůj cíl. Pokud měl vytvořit teorii, potřeboval si přisvojit slovo *informace*. Napsal: „Přestože zde uvedená ‚informace' souvisí s běžně používaným významem tohoto slova, neměli bychom ji s ním zaměňovat." Stejně jako Nyquist a Hartley, i on si přál ponechat stranou „psychologické faktory" a chtěl se soustředit pouze na ty „fyzické". Když ale informaci oddělíme od sémantického kontextu, co z ní potom zůstane? Můžeme ji nyní definovat několika způsoby, a na první pohled všechny vypadají paradoxně - ve zkratce tak můžeme říct, že informace je neurčitost, překvapení, obtížnost nebo entropie.

- „Informace úzce souvisí s neurčitostí." Neurčitost se dá měřit tak, že se spočítají všechny možné zprávy. Pokud je možná pouze jedna zpráva, není zde žádná neurčitost, a tedy ani informace.

- Některé zprávy mohou být pravděpodobnější než jiné a informace s sebou nese různou míru překvapení. Překvapení je způsob, jak mluvit o pravděpodobnostech. Pokud za písmenem *t* (v angličtině) následuje *h*, nepřenáší se nijak obsáhlá informace, protože pravděpodobnost písmene *h* byla celkem vysoká; je-li tam něco jiného, budeme překvapeni více.
- „Důležitá je obtížnost přenosu zprávy z jednoho místa na druhé." Možná se to zdálo jako obrácené či tautologické, něco jako definovat hmotnost na základě síly potřebné k uvedení nějakého tělesa do pohybu. Hmotnost tak ale *lze* definovat.
- Informace je entropie. Toto byl nejzvláštnější a nejpůsobivější nápad. Entropie, která je sama o sobě obtížným a ne zcela jasným konceptem, je v termodynamice, vědě o teple a energii, mírou neuspořádanosti.

Ústředí Bell Labs na West Street, kterým projížděly vlaky linky High Line

Vedle řízení palby a šifrování se Shannon po celou válku zabýval hlavně těmito mlhavými představami. V bytě v Greenwich Village žil sám a s kolegy se stýkal jen zřídka. Ti nyní pracovali většinou na centrále v New Jersey, zatímco Shannon dával přednost staré budově na West Street. Nikomu se nemusel zodpovídat. Jeho práce pro vojenský výzkum mu zajistila odklad vojenské služby, který platil i po skončení války. Bell Labs byly striktně mužskou záležitostí, ale za války potřebovala zejména výpočetní sekce mnoho schopných lidí, a tak začala najímat ženy. Mezi nimi byla Betty Moreová, která vyrůstala na Staten Islandu. Zpočátku pro ni byla nová práce jen něco jako písárna, snad jen že se zde vyžadovala znalost matematiky. Po roce Betty postoupila výš - do sekce pro výzkum mikrovln. Pracovala v budově dřívější „továrny na sušenky" Nabisco, která byla naproti hlavní budově přes West Street. Sekce v prvním patře navrhovala elektronky a v přízemí je vyráběla. Claude sem občas zašel na návštěvu. V roce 1948 začal s Betty chodit a začátkem roku 1949 se vzali. V té době už byl vědcem, o kterém všichni mluvili.

Knihoven, jež by odebíraly *The Bell System Technical Journal*, nebylo mnoho, a tak se badatelé dozvídali o „Matematické teorii komunikace" tradičním způsobem - pověděli si o ní - a tradičním způsobem ji i získali - řekli si autorovi o separát. Mnoho vědců používalo na takové žádosti předtištěné korespondenční lístky, které pak ve stále rostoucím počtu přicházely celý další rok. Ne každý Shannonovu pojednání rozuměl. Matematika byla pro mnohé techniky příliš obtížná a matematikům zase chyběl kontext, jenž byl technikům vlastní. Warren Weaver, který vedl obor přírodních věd za Rockefellerovu nadaci v horní části města, však již sdělil svému řediteli, že Shannon udělal pro teorii komunikace totéž „co Gibbs pro fyzikální chemii".[19] Weaver za války vedl vládní výzkum aplikované matematiky. Dohlížel na projekt řízení palby i na rodící se obor elektronických počítacích strojů. V roce 1949 napsal uznalou a nepříliš technickou esej o Shannonově teorii pro *Scientific American* a ještě téhož roku vyšla esej společně se Shannonovou prací v knize *The Mathematical Theory of Communication*. Na technika Bell Labs Johna Robinsona Pierceho, který byl současně svědkem zrodu tranzistoru a Shannonovy práce, zapůsobila kniha „jako bomba, možná jako bomba se zpožďovačem".[20]

Zatímco laik by řekl, že základním problémem komunikace je předat význam - být pochopen, Shannon to viděl jinak:

> Základní problém komunikace spočívá v tom, že v jednom bodě má více-méně přesně reprodukovat zprávu vybranou v jiném bodě.[21]

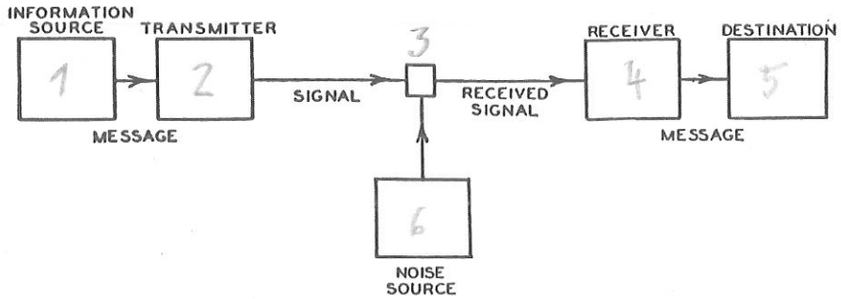
Slovo *bod* zvolil Shannon záměrně - původce a příjemce (zdroj a cíl) zprávy mohou být odděleny v čase i v prostoru. Za komunikaci lze totiž považovat

i uchovávání informací, třebas na gramofonové desce. Rovněž platí, že se zpráva netvoří, ale *vybírá* – je to volba. Může to být hodnota karty sejmuté z balíčku, tři desítkové číslice vybrané z tisíce možností nebo kombinace slov z daného seznamu kódů. Shannon mohl stěží ignorovat význam úplně, a tak ho nejprve vybavil vědeckou definicí a potom vykázal za dveře:

> Zprávy často mívají *význam* – odkazují nebo se vztahují k nějakému systému s určitými fyzickými nebo konceptuálními entitami. Tyto sémantické aspekty komunikace k technickému problému nepatří.

Weaver považoval za nutné podrobně vyložit, že tato definice pojem komunikace nezužuje, naopak že ho rozšiřuje, takže nyní zahrnuje vše: „nejen psanou a mluvenou řeč, ale také hudbu, malířství, divadlo, balet a vlastně veškeré lidské chování". A nejen lidské. Proč by stroje neměly mít zprávy, které chtějí poslat?

Shannon vyjádřil svůj model komunikace jednoduchým schématem – a nebyla náhoda, že dost podobným, jaké použil ve své tajné zprávě o kryptografii:



Podle Shannona je nutné, aby komunikační systém obsahoval tyto prvky:

1. • Informační zdroj: je jím člověk nebo stroj, který vytváří zprávu. Zprávou může být posloupnost znaků, jako u telegrafu nebo dálnopisu, nebo se dá vyjádřit matematicky jako funkce $f(x, y, t)$ času a dalších proměnných. Ve složitém příkladu, jakým je například barevná televize, jsou to tři funkce v trojrozměrném prostoru.
2. • Vysílač „nějak zpracuje zprávu" – zakóduje ji – aby vytvořil vhodný signál. Telefon mění akustický tlak na analogový elektrický proud. Telegraf zakóduje znaky do podoby teček, čárek a mezer. Složitější zprávy lze vzorkovat, komprimovat, kvantovat a prokládat.

3. • Kanál: „médium, které se používá k přenosu signálu".
4. • Přijímač dělá totéž co vysílač, ale naopak. Dekóduje nebo rekonstruuje zprávu ze signálu.
5. • Cíl je „člověk (nebo věc)" na opačném konci.

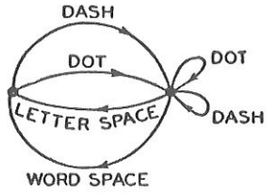
V případě mluvené řeči jsou těmito prvky mozek mluvčího, jeho hlasivky, vzduch, ucho posluchače a posluchačův mozek.

Stejně velkou roli jako ostatní prvky v Shannonově schématu hraje „zdroj
6. šumu" – technik se mu totiž nevyhne. Jedná se o vše, co poškozuje signál, ať už předvídatelně nebo ne: nechtěné příměsi, obyčejné chyby, náhodné poruchy, atmosférické poruchy, interference a zkreslení. V každém případě jsou to nepříjemné vlivy. Shannon řešil dva druhy systémů – spojité a diskrétní. V diskrétním systému mají zpráva a signál podobu individuálních samostatných symbolů, například písmen, číslic nebo teček a čárek. S výjimkou telegrafie to však byly právě spojité systémy, jejichž chybám elektrotechnici čelili každý den. Když technika požádali, aby přenosovým kanálem poslal více informací, vždy věděl, co má dělat – zvýšit výkon. Na dlouhé vzdálenosti to však nefungovalo, protože neustálé zesilování signálu vede k paralyzujícímu hromadění šumu.

Shannon se tomuto problému vyhnul tak, že se signálem zacházel jako s řetězcem samostatných, diskrétních symbolů. Místo zesilování výkonu odesilatel zdolá šum pomocí dodatečných symbolů pro opravu chyb – tak jako vzdálený posluchač pochopí afrického bubeníka nikoli díky prudšímu bubnování, ale díky jeho zvýšené mnohomluvnosti. Shannon považoval diskrétní systém za fundamentálnější i v matematickém smyslu slova. A uvažoval také o tom, že převod zpráv do diskrétní podoby se dá využít nejen při tradiční komunikaci, ale i v novém, poněkud ezoterickém oboru – teorii počítacích strojů.

A tak se vrátil k telegrafu. Přesně vzato telegraf nepoužíval jazyk s pouhými dvěma symboly, tečkou a čárkou. Ve skutečnosti telegrafisté používali tečku (jednu časovou jednotku „linka uzavřena" a jednu jednotku „linka otevřena"), čárku (například tři jednotky uzavřené linky a jednu jednotku otevřené linky) a také dvě různé mezery: mezeru mezi písmeny (tři jednotky otevřené linky) a delší mezeru, jež odděluje slova (šest jednotek otevřené linky). Tyto čtyři symboly neměly stejné postavení a pravděpodobnost. Například mezera nikdy nemohla následovat za jinou mezerou, zatímco tečka a čárka mohly následovat za čímkoli. Shannon to vyjádřil formou *stavů* – systém má podle něho dva stavy. V prvním stavu byla předchozím symbolem mezera, za níž je dovoleno použít jen tečku nebo čárku a pak systém přejde do druhého

stavu. Ve druhém stavu je dovoleno použít jakýkoli symbol a stav se změní jen po vyslání mezery. Znázornil to graficky:



To už bylo něco úplně jiného než jednoduchý binární kódovací systém. Shannon nicméně ukázal, jak odvodit korektní rovnice pro informační obsah a kapacitu kanálu. Ještě důležitější bylo, že se soustředil na statistickou strukturu jazyka dané zprávy. Již samotná existence této struktury - větší četnost *e* než *q*, *th* než *xp* a podobně - umožňuje ušetřit čas nebo kapacitu kanálu.

> V omezené míře k tomu již dochází v telegrafii, kde se používá nejkratší posloupnost přenosového kanálu - tečka - pro nejčastější anglické písmeno *E*, zatímco nepříliš častá písmena, jako *Q*, *X* a *Z*, jsou zastoupena delšími posloupnostmi teček a čárek. Tato idea se ještě rozvinula v některých obchodních kódech, kde jsou běžná slova a fráze zastoupeny čtyřpísmennými nebo pětipísmennými skupinami kódů, které v průměru ušetří značné množství času. Dnes používané telegramy s typickými blahopřáními dospěly až tak daleko, že jednu nebo dvě věty zakódují do poměrně krátké posloupnosti čísel.[22]

Aby Shannon osvětlil strukturu zprávy, uchýlil se k metodologii a jazyku fyziky stochastických procesů, od Brownova pohybu po stelární dynamiku. (Citoval zásadní pojednání Subrahmanyana Chandrasekhara, které vyšlo roku 1943 v časopise *Reviews of Modern Physics*.)[23] Stochastický proces není ani zcela deterministický (kdy bychom mohli s jistotou určit další děj), ani zcela náhodný (kdy dalším dějem může být cokoli). O pokračování rozhoduje množina pravděpodobností: každý děj má určitou pravděpodobnost, která závisí na stavu systému a zřejmě i na jeho minulosti. Pokud děj nahradíme symbolem, pak přirozený psaný jazyk, jako angličtina nebo čínština, představuje stochastický proces. Stejně tak je stochastickým procesem digitalizovaná řeč nebo televizní signál.

Když se do toho ponoříme hlouběji, Shannon prozkoumal statistickou strukturu z hlediska otázky, jak velká část zprávy ovlivňuje pravděpodobnost dalšího symbolu. Dalo by se říci, že žádná - každý symbol má svou pravděpodobnost, ale nezávisí na tom, co bylo před ním. To je případ prvního řádu. Případ druhého řádu vypadá tak, že pravděpodobnost každého symbolu závisí na bezprostředně předcházejícím symbolu, ale nikoli na ostatních. Pak má svou pravděpodobnost každé spojení dvou znaků, digram - v angličtině má například *th* větší pravděpodobnost než *xp*. V případě třetího řádu sledujeme třípísmenná spojení a podobně. V běžném textu je kromě toho rozumné sledovat úroveň slov místo jednotlivých písmen a do hry vstupuje i mnoho různých statistických závislostí. Například ihned po slově *žlutý* následují některá slova s mnohem vyšší pravděpodobností, než je běžné, zatímco pravděpodobnost jiných se blíží nule. Anglická slova po členu *an*, která by začínala souhláskami, jsou nesmírně vzácná. Pokud je písmeno *u* na konci nějakého slova, je tím slovem nejspíše *you*. Jestliže jsou dvě písmena za sebou stejná, jsou to pravděpodobně *ll, ee, ss* nebo *oo*. Vnitřní uspořádání textu může působit i na velké vzdálenosti - ve zprávě, která obsahuje slovo *kráva*, je velmi pravděpodobné, že se toto slovo objeví znovu, i když mezi nimi bude mnoho dalších písmen. Stejné je to se slovem *kůň*. Shannon ukázal, že zpráva se může chovat jako dynamický systém, jehož budoucí směřování ovlivňuje jeho minulost.

Na ilustraci rozdílů mezi jednotlivými úrovněmi struktury vytvořil - vypočítal - řadu „aproximací" anglického textu. Použil abecedu s 27 znaky, jimiž byla písmena plus mezera mezi slovy, a s pomocí tabulky náhodných čísel vytvářel řetězce znaků. (Náhodná čísla čerpal z nové knihy, kterou k těmto účelům vydalo nakladatelství Cambridge University Press. Jejích 100 000 číslic stálo tři šilinky a devět pencí a autoři poskytli „záruku náhodného uspořádání".)[24] I přes použití předpřipravených náhodných čísel byla tvorba vět nesmírně pracná. Výsledky vypadaly následovně:

- „Aproximace nultého řádu" - tedy pouze náhodné znaky, žádná struktura ani korelace.
  XFOML RXKHRJFFJUJ ZLPWCFWKCYJ
  FFJEYVKCQSGHYD GPAAMKBZAACIBZLHJGD.

- První řád - navzájem nezávislá písmena, ale jejich četnost je taková, jaká se v angličtině očekává: více písmen *e* a *t*, méně *z* a *j*, délka slov vypadá přiměřeně.
  OCRO HLI RGWR NIMILWIS EU LL NBNESEBYA TH EEI ALHENHTTPA
  OOBTTVA NAH BRL.

- Druhý řád - četnost každého písmene odpovídá angličtině, stejně jako četnost každého digramu - skupiny dvou písmen. (Shannon našel nezbytné

statistiky v tabulkách pro luštitele šifer.[25] Nejběžnějším digramem v angličtině je *th*, jehož četnost je 168 na 1 000 slov. Za ním následují *he, an, re* a *er*, nemálo digramů má nulovou četnost.)
ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO TIZIN ANDY TOBESEACE CTISBE.

- Třetí řád - struktura zohledňuje trigramy.
  IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE.

- Aproximace prvního řádu na úrovni slov.
  REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT NATURAL HERE HE THE A IN CAME THE TO OF TO EXPERT GRAY COME TO FURNISHES THE LINE MESSAGE HAD BE THESE.

- Aproximace prvního řádu na úrovni slov - nyní mají očekávanou četnost dvojice slov, takže neuvidíme spojení jako „a in" nebo „to of".
  THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED.

Tyto věty postupně stále více „vypadají" jako angličtina. To bylo změřeno i objektivněji: lidé, kteří zvládají psaní deseti prsty, píší tyto věty rychleji. To je další doklad, jak si nevědomky osvojujeme statistickou strukturu jazyka.

Kdyby měl Shannon víc času, mohl vytvořit další aproximace, ale práce na nich již začínala být neúnosná. Cílem bylo ukázat zprávu jako výsledek procesu, který generoval události s diskrétními pravděpodobnostmi. Co tedy uvést k množství informace či k rychlosti, se kterou informace vzniká? Pro každou událost platí, že možné volby mají známé pravděpodobnosti (označíme je p_1, p_2, p_3 atd.). Shannon chtěl definovat míru informace (označíme ji H) jako míru neurčitosti: „[míru] toho, kolik ‚volby' s sebou nese výběr události či jak nejistí jsme v otázce výsledku."[26] Tyto pravděpodobnosti mohou být různé, ale obecně platí, že více možností na výběr znamená více neurčitosti, tedy větší množství informace. Možnost volby většího celku lze rozčlenit na posloupnost voleb jeho složek (s jejich vlastními pravděpodobnostmi), a tyto pravděpodobnosti se musejí chovat aditivně - například pravděpodobnost konkrétního digramu musí být váženým součtem pravděpodobností jednotlivých symbolů. Kdyby

byly tyto pravděpodobnosti stejné, pak množství informace, které přenáší každý symbol, bude logaritmem počtu možných symbolů - tedy Nyquistův a Hartleyův vzorec:

$$H = n \log s$$

Pro realističtější případ předvedl Shannon elegantní řešení problému, jak měřit informaci jako funkci pravděpodobností - rovnici, která sčítala pravděpodobnosti s logaritmickým vahami (nejvíce vyhovoval základ 2). Je to průměrný logaritmus nepravděpodobnosti zprávy, vlastně míra neočekávanosti:

$$H = -\Sigma\, p_i \log_2 p_i$$

kde p_i je pravděpodobnost každé zprávy. Shannon prohlásil, že se s veličinou tohoto typu budeme opakovaně setkávat, neboť „hraje ústřední roli v teorii informace jako míra informace, volby a neurčitosti". H se vyskytuje opravdu všude. Běžně se mu říká entropie zprávy či Shannonova entropie nebo zkrátka informace.

Bylo třeba zavést novou jednotku. Shannon uvedl: „Výsledné jednotky se mohou nazývat binární číslice nebo zkráceně *bity*."[27] Bit je nejmenší možné množství informace a jako takový představuje množství neurčitosti, které existuje při házení mincí. Když si hodíme mincí, vytvoříme volbu mezi dvěma stejně pravděpodobnými možnostmi - v tomto případě se p_1 i p_2 rovnají 1/2 a protože $\log_2 1/2 = -1$, platí $H = 1$ bit. Jediné písmeno, které je náhodně vybrané z abecedy s 32 písmeny, předává větší množství informace - konkrétně pět bitů, protože existuje 32 možných zpráv, $p_i = 1/32$ a $\log_2 1/32 = -5$. Řetězec 1 000 takových písmen přenáší 5 000 bitů (vyjde to i prostým vynásobením, ale raději si ukážeme korektní odvození), neboť množství informace je množství neurčitosti, čili počet možných voleb, a 1 000 znaků v abecedě s 32 písmeny znamená, že existuje 32^{1000} možných zpráv a $\log_2 32^{1000} = 1\,000 \log_2 32 = 5\,000$.

Zde opět vstupuje na scénu statistická struktura přirozených jazyků. Pokud je text s 1 000 znaky v angličtině, počet možných zpráv je menší, a to *značně*. Když vezmeme ohled na korelace do vzdálenosti 8 znaků, má angličtina podle Shannonova odhadu redundanci asi 50 % - každý znak zprávy nenese 5 bitů, ale jen zhruba 2,3. Pokud zohledníme vztahy na ještě větší vzdálenosti - na úrovni vět a odstavců - zvýší se podle něj tento odhad až na 75 %. Zde však varoval, že takové odhady jsou „velmi nejisté a nespolehlivé, neboť silně závisejí na druhu konkrétního textu".[28] Jeden ze způsobů, jak měřit redundan-

ci, byl velmi empirický a spočíval v psychologickém testu. Tato metoda „využívá skutečnosti, že každý, kdo mluví nějakým jazykem, má sám od sebe obrovské znalosti statistických vlastností tohoto jazyka".

> Znalost slov, idiomů, klišé a gramatiky mu při čtení textu umožňuje doplnit chybějící písmena, změnit ta nesprávná či doplnit v rozhovoru nedokončenou frázi.

Místo „mu" také mohl říci „jí", protože pokusy prováděl na své ženě Betty. Vzal z police jednu knihu (byla to detektivka *Výpalné na Noon Street* od Raymonda Chandlera), náhodně prstem zakryl krátký úryvek a požádal Betty, aby zkusila uhodnout jedno písmeno, pak druhé, třetí a další. Čím více textu viděla, tím větší měla šanci na správnou odpověď. Po slovech „*a small oblong reading lamp on the*" („malá podlouhlá lampička na") tipovala další písmeno špatně, jakmile se však dozvěděla, že je to *d*, další tři písmena dokázala uhodnout snadno. Shannon poznamenal: „Chyby se podle očekávání nejčastěji objevují na začátku slov a slabik, kde má směr myšlení více příležitostí odbočit jinam."

Takovéto měření předvídatelnosti a redundance je vlastně opakem měření informačního obsahu. Pokud se dá písmeno odhadnout podle toho, co je před ním, je nadbytečné, a co je nadbytečné, nenese žádnou informaci. Pokud je redundance v angličtině 75 %, pak anglicky psaná zpráva s 1 000 písmeny předává jen 25 % informace proti 1 000 náhodně zvoleným písmenům. Možná to zní paradoxně, ale náhodné zprávy přenášejí *větší* množství informace, z čehož plyne, že text v přirozeném jazyce se dá pro přenos nebo uložení kódovat účinněji než text náhodný.

Shannon předvedl jeden způsob, jak to udělat - algoritmus, který využívá různých pravděpodobností jednotlivých symbolů - a dodal k tomu mnoho zajímavých výsledků. Jedním z nich byl vzorec pro výpočet kapacity přenosového kanálu čili absolutní hranice rychlosti jakéhokoli komunikačního kanálu (dnes jednodušeji nazývaná Shannonův limit). Dalším byl objev, že v rámci tohoto limitu vždy lze navrhnout takové schéma opravy chyb, jež umožní překonat jakoukoli úroveň šumu. Odesilatel může být nucen věnovat opravě chyb víc a víc bitů a přenos bude stále pomalejší, ale zpráva se nakonec k příjemci dostane. Shannon nepředvedl, jak konkrétně taková schémata navrhovat, pouze dokázal, že to je možné - poskytl tím téma pro budoucí badatele v informatice. Jeho kolega Robert Fano o řadu let později vzpomínal: „Snížit pravděpodobnost chyby na tak malé číslo, jak si přejete? Nikoho taková možnost nenapadla. Nevím, jak na to přišel a co ho přivedlo k závěru, že by to mohlo fungovat.

Nicméně na této práci se zakládá téměř celá moderní teorie komunikace."[29] Ať už se redundance odstraňuje s cílem zvýšit účinnost nebo se přidává s cílem umožnit opravu chyb, kódování závisí na znalosti statistické struktury jazyka. Informaci nelze oddělit od pravděpodobností. Bit je v podstatě vždy hozením mincí.

Pokud jsou dvě strany mince jedním způsobem, jak znázornit bit, Shannon předvedl i praktičtější příklad:

> Zařízení se dvěma stálými stavy, jako je relé nebo klopný obvod, může uložit jeden bit informace. *N* takových zařízení dokáže uložit *N* bitů, jelikož celkový počet možných stavů je 2^N a $\log_2 2^N = N$.

Shannon viděl přístroje - třebas reléová pole - které dokázaly uložit stovky i tisíce bitů. Tehdy to vypadalo jako opravdu velké množství. V době, kdy dokončoval svoji práci, zašel do kanceláře Williama Shockleyho. Tento fyzik, třicátník, byl jeho kolega z Bell Labs a patřil k sekci fyziky pevných látek. Tito vědci pracovali na elektronických alternativách elektronek. Na stole měl maličký model, polovodičový krystal. Řekl Shannonovi: „Tohle je polovodičový zesilovač."[30] V té chvíli se ještě nijak nejmenoval.

V létě roku 1949, před vydáním *The Mathematical Theory of Communication*, si Shannon vzal tužku a papír, načrtl čáru odshora dolů (viz obrázek na další straně) a vypsal exponenty od 10^0 do 10^{13}. Osu popsal slovy „úložná kapacita v bitech",[31] a začal si na ni vynášet položky, o kterých by se dalo říci, že „uchovávají" informaci. Například číselník telefonu (10 desítkových číslic) má informační kapacitu něco přes tři bity. U 10^3 bitů Shannon napsal: „děrný štítek (všechny konfigurace dovoleny)." U 10^4 uvedl: „strana s jednoduchým řádkováním (32 možných symbolů)." Vedle 10^5 napsal něco zcela nekonvenčního: „genetická výbava člověka." To zatím žádného vědce nenapadlo - Jamesi D. Watsonovi, který studoval zoologii v Indianě, bylo jednadvacet let a do jeho objevu struktury DNA ještě pár let chybělo. Poprvé v historii teď někdo tvrdil, že genom je informační úložiště a že dá se měřit v bitech. Shannonův odhad byl hodně opatrný, alespoň o čtyři řády. Domníval se, že „gramofonová deska (128 úrovní amplitudy)" má větší informační obsah - asi 300 000 bitů. K úrovni 10^7 bitů přiřadil tlustý odborný časopis (*Proceedings of the Institute of Radio Engineers*) a k úrovni 10^9 připsal *Encyclopaedia Britannica*. Jednu hodinu televizního vysílání odhadl na 10^{11} bitů a hodinu „barevného filmu" na více než 10^{12} bitů. Těsně pod poslední značkou 10^{14}, 100 bilionů bitů, připsal největší informační úložiště, které ho napadlo - Knihovnu Kongresu.