Mnemonic code for generating *deterministic keys*

Marek Palatinus · Pavol Rusnak · Aaron Voisine · Sean Bowe

slush@satoshilabs.com · stick@satoshilabs.com · voisine@gmail.com · ewillbefull@gmail.com

39

Abstract

This BIP describes the implementation of a mnemonic code or mnemonic sentence -- a group of easy to remember words -- for the generation of deterministic wallets.

It consists of two parts: generating the mnemonic and converting it into a binary seed. This seed can be later used to generate deterministic wallets using BIP-0032 or similar methods.

Copyright

This BIP falls under the MIT License.

Motivation

A mnemonic code or sentence is superior for human interaction compared to the handling of raw binary or hexadecimal representations of a wallet seed. The sentence could be written on paper or spoken over the telephone.

This guide is meant to be a way to transport computer-generated randomness with a human-readable transcription. It's not a way to process user-created sentences (also known as brain-wallets) into a wallet seed.

Generating the mnemonic

The mnemonic must encode entropy in a multiple of 32 bits. With more entropy security is improved but the sentence length increases. We refer to the initial entropy length as ENT. The allowed size of ENT is 128-256 bits.

First, an initial entropy of ENT bits is generated. A checksum is generated by taking the first `ENT / 32` bits of its SHA256 hash. This checksum is appended to the end of the initial entropy. Next, these concatenated bits are split into groups of 11 bits, each encoding a number from 0-2047, serving as an index into a wordlist. Finally, we convert these numbers into words and use the joined words as a mnemonic sentence.

The following table describes the relation between the initial entropy length (ENT), the checksum length (CS), and the length of the generated mnemonic sentence (MS) in words.

```
CS = ENT / 32
MS = (ENT + CS) / 11

|  ENT  | CS | ENT+CS |  MS  |
+-------+----+--------+------+
|  128  |  4 |   132  |  12  |
|  160  |  5 |   165  |  15  |
|  192  |  6 |   198  |  18  |
|  224  |  7 |   231  |  21  |
|  256  |  8 |   264  |  24  |
```

Wordlist

An ideal wordlist has the following characteristics:

a) smart selection of words

- the wordlist is created in such a way that it's enough to type the first four letters to unambiguously identify the word

b) similar words avoided

- word pairs like "build" and "built", "woman" and "women", or "quick" and "quickly" not only make remembering the sentence difficult but are also more error prone and more difficult to guess

c) sorted wordlists

- the wordlist is sorted which allows for more efficient lookup of the code words (i.e. implementations can use binary search instead of linear search)
- this also allows trie (a prefix tree) to be used, e.g. for better compression

The wordlist can contain native characters, but they must be encoded in UTF-8 using Normalization Form Compatibility Decomposition (NFKD).

From mnemonic to seed

A user may decide to protect their mnemonic with a passphrase. If a passphrase is not present, an empty string "" is used instead.

To create a binary seed from the mnemonic, we use the PBKDF2 function with a mnemonic sentence (in UTF-8 NFKD) used as the password and the string "mnemonic" + passphrase (again in UTF-8 NFKD) used as the salt. The iteration count is set to 2048 and HMAC-SHA512 is used as the pseudo-random function. The length of the derived key is 512 bits (= 64 bytes).

This seed can be later used to generate deterministic wallets using BIP-0032 or similar methods.

The conversion of the mnemonic sentence to a binary seed is completely independent from generating the sentence. This results in a rather simple code; there are no constraints on sentence structure and clients are free to implement their own wordlists or even whole sentence generators, allowing for flexibility in wordlists for typo detection or other purposes.

Although using a mnemonic not generated by the algorithm described in "Generating the mnemonic" section is possible, this is not advised and software must compute a checksum for the mnemonic sentence using a wordlist and issue a warning if it is invalid.

The described method also provides plausible deniability, because every passphrase generates a valid seed (and thus a deterministic wallet) but only the correct one will make the desired wallet available.

Wordlists

Since the vast majority of BIP39 wallets supports only the English wordlist, it is **strongly discouraged** to use non-English wordlists for generating the mnemonic sentences.

If you still feel your application really needs to use a localized wordlist, use one of the following instead of inventing your own.

- Wordlists

Test vectors

The test vectors include input entropy, mnemonic and seed. The passphrase "TREZOR" is used for all vectors.

`https://github.com/trezor/python-mnemonic/blob/master/vectors.json`

Also see `https://github.com/bip32JP/bip32JP.github.io/blob/master/test_JP_BIP39.json`

(Japanese wordlist test with heavily normalized symbols as passphrase)

Shortcomings

Some shortcomings have been identified with this proposal:

- Generated seed depends on the wordlist that was used for the mnemonic. Because the "mnemonic to seed" process uses the mnemonic sentence directly rather than the original entropy, translating the mnemonic to a different wordlist necessarily creates a completely different seed. This is not an issue if you only support the English wordlist, as recommended above.
- Because the seed is generated by hashing the mnemonic, it is not possible to represent an arbitrary BIP-0032 seed via a BIP-0039 sentence: the conversion is one-way only (from BIP-0039 sentence to BIP-0032 seed).
- The checksum is short. This means it only gives modest odds of catching random errors (1-in-256 errors will be missed). It is also not able to provide any assistance in correcting errors.
- No versioning scheme. When originally introduced, there was no way to distinguish the address format that should be used for a BIP-0039 key. This is now largely mitigated by use of descriptor wallets (BIP-0380) in addition to a seed however.

Related Work

The authors of BIP-0039 proposed the SLIP-0039 scheme as an intended successor to BIP-0039 improving on the above shortcomings.

Reference Implementation

Reference implementation including wordlists is available from

`http://github.com/trezor/python-mnemonic`

bip-0039/english.txt

abandon apart behind burst cinnamon crew diesel elephant faith fox grief ignore kitchen luxury moon oil piano purity review security snack struggle thunder type voyage
ability apology believe bus circle cricket diet elevator fall fragile grit ill kite lyrics moral okay picnic purpose reward seed snake student ticket typical wage
able appear below business citizen crime differ elite false frame grocery illegal kitten machine more old picture purse rhythm seek snap stuff tide ugly wagon
about apple belt busy city crisp digital else fame frequent group illness kiwi mad morning olive piece push rib segment sniff stumble tiger umbrella wait
above approve bench butter civil critic dignity embark family fresh grow image knee magic mosquito olympic pig put ribbon select snow style tilt unable walk
absent april benefit buyer claim crop dilemma embody famous friend grunt imitate knife magnet mother omit pigeon puzzle rice sell soap subject timber unaware wall
absorb arch best buzz clap cross dinner embrace fan fringe guard immense knock maid motion once pill pyramid rich seminar soccer submit time uncle walnut
abstract arctic betray cabbage clarify crouch dinosaur emerge fancy frog guess immune know mail motor one pilot quality ride senior social subway tiny uncover want
absurd area better cabin claw crowd direct emotion fantasy front guide impact lab main mountain onion pink quantum ridge sense sock success tip under warfare
abuse arena between cable clay crucial dirt employ farm frost guilt impose label major mouse online pioneer quarter rifle sentence soda such tired undo warm
access argue beyond cactus clean cruel disagree empower fashion frown guitar improve labor make move only pipe question right series soft sudden tissue unfair warrior
accident arm bicycle cage clerk cruise discover empty fat frozen gun impulse ladder mammal movie open pistol quick rigid service solar suffer title unfold wash
account armed bid cake clever crumble disease enable fatal fruit gym inch lady man much opera pitch quit ring session soldier sugar toast unhappy wasp
accuse armor bike call click crunch dish enact father fuel habit include lake manage muffin opinion pizza quiz riot settle solid suggest tobacco uniform waste
achieve army bind calm client crush dismiss end fatigue fun hair income lamp mandate mule oppose place quote ripple setup solution suit today unique water
acid around biology camera cliff cry disorder endless fault funny half increase language mango multiply option planet rabbit risk seven solve summer toddler unit wave
acoustic arrange bird camp climb crystal display endorse favorite furnace hammer index laptop mansion muscle orange plastic raccoon ritual shadow someone sun toe universe way
acquire arrest birth can clinic cube distance enemy feature fury hamster indicate large manual museum orbit plate race rival shaft song sunny together unknown wealth
across arrive bitter canal clip culture divert energy february future hand indoor later maple mushroom orchard play rack river shallow soon sunset toilet unlock weapon
act arrow black cancel clock cup divide enforce federal gadget happy industry latin marble music order please radar road share sorry super token until wear
action art blade candy clog cupboard divorce engage fee gain harbor infant laugh march must ordinary pledge radio roast shed sort supply tomato unusual weasel
actor artefact blame cannon close curious dizzy engine feed galaxy hard inflict laundry margin mutual organ pluck rail robot shell soul supreme tomorrow unveil weather
actress artist blanket canoe cloth current doctor enhance feel gallery harsh inform lava marine myself orient plug rain robust sheriff sound sure tone update web
actual artwork blast canvas cloud curtain document enjoy female game harvest inhale law market mystery original plunge raise rocket shield soup surface tongue upgrade wedding
adapt ask bleak canyon clown curve dog enlist fence gap hat inherit lawn marriage myth orphan poem rally romance shift source surge tonight uphold weekend
add aspect bless capable club cushion doll enough festival garage have initial lawsuit mask naive ostrich poet ramp roof shine south surprise tool upon weird
addict assault blind capital clump custom dolphin enrich fetch garbage hawk inject layer mass name other point ranch rookie ship space surround tooth upper welcome
address asset blood captain cluster cute domain enroll fever garden hazard injury lazy master napkin outdoor polar random room shiver spare survey top upset west
adjust assist blossom car clutch cycle donate ensure few garlic head inmate leader match narrow outer pole range rose shock spatial suspect topic urban wet
admit assume blouse carbon coach dad donkey enter fiber garment health inner leaf material nasty output police rapid rotate shoe spawn sustain topple urge whale
adult asthma blue card coast damage donor entire fiction gas heart innocent learn math nation outside pond rare rough shoot speak swallow torch usage what
advance athlete blur cargo coconut damp door entry field gasp heavy input leave matrix nature oval pony rate round shop special swamp tornado use wheat
advice atom blush carpet code dance dose envelope figure gate hedgehog inquiry lecture matter near oven pool rather route short speed swap tortoise used wheel
aerobic attack board carry coffee danger double episode file gather height insane left maximum neck over popular raven royal shoulder spell swarm toss useful when
affair attend boat cart coil daring dove equal film gauge hello insect leg maze need own portion raw rubber shove spend swear total useless where
afford attitude body case coin dash draft equip filter gaze helmet inside legal meadow negative owner position razor rude shrimp sphere sweet tourist usual whip
afraid attract boil cash collect daughter dragon era final general help inspire legend mean neglect oxygen possible ready rug shrug spice swift toward utility whisper
again auction bomb casino color dawn drama erase find genius hen install leisure measure neither oyster post real rule shuffle spider swim tower vacant wide
age audit bone castle column day drastic erode fine genre hero intact lemon meat nephew ozone potato reason run shy spike swing town vacuum width
agent august bonus casual combine deal draw erosion finger gentle hidden interest lend mechanic nerve pact pottery rebel runway sibling spin switch toy vague wife
agree aunt book cat come debate dream error finish genuine high into length medal nest paddle poverty rebuild rural sick spirit sword track valid wild
ahead author boost catalog comfort debris dress erupt fire gesture hill invest lens media net page powder recall sad side split symbol trade valley will
aim auto border catch comic decade drift escape firm ghost hint invite leopard melody network pair power receive saddle siege spoil symptom traffic valve win
air autumn boring category common december drill essay first giant hip involve lesson melt neutral palace practice recipe sadness sight sponsor syrup tragic van window
airport average borrow cattle company decide drink essence fiscal gift hire iron letter member never palm praise record safe sign spoon system train vanish wine
aisle avocado boss caught concert decline drip estate fish giggle history island level memory news panda predict recycle sail silent sport table transfer vapor wing
alarm avoid bottom cause conduct decorate drive eternal fit ginger hobby isolate liar mention next panel prefer reduce salad silk spot tackle trap various wink
album awake bounce caution confirm decrease drop ethics fitness giraffe hockey issue liberty menu nice panic prepare reflect salmon silly spray tag trash vast winner
alcohol aware box cave congress deer drum evidence fix girl hold item library mercy night panther present reform salon silver spread tail travel vault winter
alert away boy ceiling connect defense dry evil flag give hole ivory license merge noble paper pretty refuse salt similar spring talent tray vehicle wire
alien awesome bracket celery consider define duck evoke flame glad holiday jacket life merit noise parade prevent region salute simple spy talk treat velvet wisdom
all awful brain cement control defy dumb evolve flash glance hollow jaguar lift merry nominee parent price regret same since square tank tree vendor wise
alley awkward brand census convince degree dune exact flat glare home jar light mesh noodle park pride regular sample sing squeeze tape trend venture wish
allow axis brass century cook delay during example flavor glass honey jazz like message normal parrot primary reject sand siren squirrel target trial venue witness
almost baby brave cereal cool deliver dust excess flee glide hood jealous limb metal north party print relax satisfy sister stable task tribe verb wolf
alone bachelor bread certain copper demand dutch exchange flight glimpse hope jeans limit method nose pass priority release satoshi situate stadium taste trick verify woman
alpha bacon breeze chair copy demise duty excite flip globe horn jelly link middle notable patch prison relief sauce six staff tattoo trigger version wonder
already badge brick chalk coral denial dwarf exclude float gloom horror jewel lion midnight note path private rely sausage size stage taxi trim very wood
also bag bridge champion core dentist dynamic excuse flock glory horse job liquid milk nothing patient prize remain save skate stairs teach trip vessel wool
alter balance brief change corn deny eager execute floor glove hospital join list million notice patrol problem remember say sketch stamp team trophy veteran word
always balcony bright chaos correct depart eagle exercise flower glow host joke little mimic novel pattern process remind scale ski stand tell trouble viable work
amateur ball bring chapter cost depend early exhaust fluid glue hotel journey live mind now pause produce remove scan skill start ten truck vibrant world
amazing bamboo brisk charge cotton deposit earn exhibit flush goat hour joy lizard minimum nuclear pave profit render scare skin state tenant true vicious worry
among banana broccoli chase couch depth earth exile fly goddess hover judge load minor number payment program renew scatter skirt stay tennis truly victory worth
amount banner broken chat country deputy easily exist foam gold hub juice loan minute nurse peace project rent scene skull steak tent trumpet video wrap
amused bar bronze cheap couple derive east exit focus good huge jump lobster miracle nut peanut promote reopen scheme slab steel term trust view wreck
analyst barely broom check course describe easy exotic fog goose human jungle local mirror oak pear proof repair school slam stem test truth village wrestle
anchor bargain brother cheese cousin desert echo expand foil gorilla humble junior lock misery obey peasant property repeat science sleep step text try vintage wrist
ancient barrel brown chef cover design ecology expect fold gospel humor junk logic miss object pelican prosper replace scissors slender stereo thank tube violin write
anger base brush cherry coyote desk economy expire follow gossip hundred just lonely mistake oblige pen protect report scorpion slice stick that tuition virtual wrong
angle basic bubble chest crack despair edge explain food govern hungry kangaroo long mix obscure penalty proud require scout slide still theme tumble virus yard
angry basket buddy chicken cradle destroy edit expose foot gown hunt keen loop mixed observe pencil provide rescue scrap slight sting then tuna visa year
animal battle budget chief craft detail educate express force grab hurdle keep lottery mixture obtain people public resemble screen slim stock theory tunnel visit yellow
ankle beach buffalo child cram detect effort extend forest grace hurry ketchup loud mobile obvious pepper pudding resist script slogan stomach there turkey visual you
announce bean build chimney crane develop egg extra forget grain hurt key lounge model occur perfect pull resource scrub slot stone they turn vital young
annual beauty bulb choice crash device eight eye fork grant husband kick love modify ocean permit pulp response sea slow stool thing turtle vivid youth
another because bulk choose crater devote either eyebrow fortune grape hybrid kid loyal mom october person pulse result search slush story this twelve vocal zebra
answer become bullet chronic crawl diagram elbow fabric forum grass ice kidney lucky moment odor pet pumpkin retire season small stove thought twenty voice zero
antenna beef bundle chuckle crazy dial elder face forward gravity icon kind luggage monitor off phone punch retreat seat smart strategy three twice void zone
antique before bunker chunk cream diamond electric faculty fossil great idea kingdom lumber monkey offer photo pupil return second smile street thrive twin volcano zoo
anxiety begin burden churn credit diary elegant fade foster green identify kiss lunar monster office phrase puppy reunion secret smoke strike throw twist volume
any behave burger cigar creek dice element faint found grid idle kit lunch month often physical purchase reveal section smooth strong thumb two vote

MIT License. Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions: The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software. THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.