

Email-parsing

Problem ID: email

Nyckelorden **MÅSTE**, **MÅSTE INTE**, **KRÄVS**, **SKALL**, **SKALL INTE**, **BÖR**, **BÖR INTE**, **REKOMMENDERAS**, **FÅR**, **FRIVILLIGT** i denna uppgiftslydelse skall tolkas som beskriver i RFC 2119.

För att lösa denna uppgift **REKOMMENDERAS** det att läsa följande specifikationer:

- RFC 20: <https://www.ietf.org/rfc/rfc20.txt>
- RFC 1866: <https://www.ietf.org/rfc/rfc1866.txt>
- RFC 2045: <https://www.ietf.org/rfc/rfc2045.txt>
- RFC 2046: <https://www.ietf.org/rfc/rfc2046.txt>
- RFC 2119: <https://www.ietf.org/rfc/rfc2119.txt>
- RFC 2616: <https://www.ietf.org/rfc/rfc2616.txt>
- RFC 2822: <https://www.ietf.org/rfc/rfc2822.txt>
- RFC 3676: <https://www.ietf.org/rfc/rfc3676.txt>

Du kommer få email i formatet som beskrivs av RFC 2822. Din uppgift är att hitta en föredragen del i detta email, vars innehåll är ett MIME-meddelande, som specificerat i RFC 2045 och RFC 2046. Texten i denna del **SKALL** du sedan skriva ut.

Meddelandet kommer alltid vara av typen `multipart/alternative`. Den föredragna delen i meddelandet är antingen av typen `text/html`, som beskrivs i RFC 1866, eller typen `text/plain`, som beskrivs i RFC 3676. Teckenkodningen kommer att vara ASCII, som beskriven i RFC 20.

För enkelhets skull kommer varje rad att termineras med en enkel LF i stället för en CRLF som i vanlig HTTP 1.1 (beskriven i RFC 2616).

Till skillnad från en riktig email-klient **FÅR** du anta att saker och ting kommer att vara snälla. T.ex. kommer HTML-en i en `text/html`-del alltid validera.

Indata

Indata innehåller ett email, se exempel nedan.

Utdata

Du **SKALL** skriva ut den extraherade texten och inget annat. Extra nyrader i början/slutet är **FRIVILLIGT**.

Poäng

20 poäng det kommer bara finnas en enda multipart-del, och den är av typen `"text/plain"`.

30 poäng den föredragna delen kommer att vara av typ `"text/plain"`.

25 poäng den föredragna delen kommer att vara av typ `"text/plain"`, och kommer att kodas med `"Content-Transfer-Encoding: quoted-printable"`.

25 poäng den föredragna delen kommer att vara av typ `"text/html"`. Alla `<HTML-taggar>` och `&entiteter`; ska tas bort.

Endast i den tredje gruppen kommer kodningen vara `"Content-Transfer-Encoding: quoted-printable"`. I de övriga kommer den att vara `"7bit"`.

Sample Input 1

```
MIME-Version: 1.0
Received: by 10.112.236.98 with HTTP; Thu, 28 Jan 2016 10:57:40 -0800 (PST)
Date: Thu, 28 Jan 2016 19:57:40 +0100
Delivered-To: simon.lindholm10@gmail.com
Message-ID: <CAK=Tff0+3K4tXGvYh+VG-qP6Ldz+q-+jaZW2Q01dkL4fDQxGkw@mail.gmail.com>
Subject: test
From: Simon Lindholm <simon.lindholm10@gmail.com>
To: simonlindholm@fastmail.fm
Content-Type: multipart/alternative; boundary=001a11c269f42ab6b0052a697e72

--001a11c269f42ab6b0052a697e72
Content-Type: text/plain; charset=UTF-8

a test

--001a11c269f42ab6b0052a697e72--
```

Sample Output 1

```
a test
```

Sample Input 2

```
Delivered-To: simon.lindholm10@gmail.com
Received: by 10.112.155.195 with SMTP id vy3csp710980lbb;
      Thu, 28 Jan 2016 14:08:38 -0800 (PST)
Return-Path: <simonlindholm@fastmail.fm>
Received: from out3-smtp.messagingengine.com (out3-smtp.messagingengine.com. [66.111.4.27]
      by mx.google.com with ESMTPS id wb3si19310522pab.114.2016.01.28.14.08.37
      for <simon.lindholm10@gmail.com>
      (version=TLS1_2 cipher=ECDHE-RSA-AES128-GCM-SHA256 bits=128/128);
      Thu, 28 Jan 2016 14:08:37 -0800 (PST)
Received: from compute2.internal (compute2.nyi.internal [10.202.2.42])
      by mailout.nyi.internal (Postfix) with ESMTP id 965F620767
      for <simon.lindholm10@gmail.com>; Thu, 28 Jan 2016 17:08:34 -0500 (EST)
Received: from web5 ([10.202.2.215])
      by compute2.internal (MEProxy); Thu, 28 Jan 2016 17:08:34 -0500
Received: by web5.nyi.internal (Postfix, from userid 99)
      id 26696A70216; Thu, 28 Jan 2016 17:08:34 -0500 (EST)
Message-Id: <1454018914.1678085.505620754.5295DB0C@webmail.messagingengine.com>
From: Simon Lindholm <simonlindholm@fastmail.fm>
To: Simon Lindholm <simon.lindholm10@gmail.com>
MIME-Version: 1.0
Content-Transfer-Encoding: 7bit
Content-Type: multipart/alternative; boundary="_-----=_145401891416780850";
      charset="utf-8"
Subject: Re: test
Date: Thu, 28 Jan 2016 23:08:34 +0100
In-Reply-To: <CAK=Tff0+3K4tXGvYh+VG-qP6Ldz+q-+jaZW2Q01dkL4fDQxGkw@mail.gmail.com>
References: <CAK=Tff0+3K4tXGvYh+VG-qP6Ldz+q-+jaZW2Q01dkL4fDQxGkw@mail.gmail.com>

This is a multi-part message in MIME format.

--_-----=_145401891416780850
Content-Transfer-Encoding: 7bit
Content-Type: text/plain

a reply

On Thu, 28 Jan 2016, at 07:57 PM, Simon Lindholm wrote:
> a test

--_-----=_145401891416780850
Content-Transfer-Encoding: 7bit
Content-Type: text/html

<!DOCTYPE html>
<html>
<head>
<title></title>
</head>
<body><div>a reply<br></div>
<div>&nbsp;</div>
<div>On Thu, 28 Jan 2016, at 07:57 PM, Simon Lindholm wrote:<br></div>
<blockquote type="cite"><div dir="ltr">a test<br></div>
</blockquote><div>&nbsp;</div>
</body>
</html>

--_-----=_145401891416780850--
```

Sample Output 2

a reply

On Thu, 28 Jan 2016, at 07:57 PM, Simon Lindholm wrote:
a test