PERSIAN SUPPLEMENT

to the

TIDES Standard for the Annotation of Temporal Expressions

July 2008

Karine Megerdoomian

Contact:  karine@mitre.org
http://time2.mitre.org

Approved for public release; distribution unlimited.

The views, opinions, and/or findings contained in this report are those of the MITRE Corporation and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.
© 2008 The MITRE Corporation. All rights reserved.
Table of Contents

1 Introduction ........................................................................................................................................ 1

2 Determining What Kinds of Expressions to Annotate ....................................................................... 2
  2.1 Markable Expressions .................................................................................................................. 2
  2.2 Non-Markables ............................................................................................................................ 3
    2.2.1 Non-Markable Parts of Speech: Prepositions and Subordinating Conjunctions .................. 3
    2.2.2 Non-Markable Point and Duration Expressions ............................................................... 4
    2.2.3 Non-Markable Frequency Expressions .............................................................................. 6
    2.2.4 Non-Markable Proper Names ............................................................................................ 6

3 Capturing the Meaning of Temporal Expressions ............................................................................. 8
  3.1 General Notes on Persian .............................................................................................................. 8
    3.1.1 Ambiguities in Persian ......................................................................................................... 8
    3.1.2 Syntax of Noun Phrases ....................................................................................................... 8
    3.1.3 Conversational Persian ....................................................................................................... 10
  3.2 Precise Temporal Expressions ...................................................................................................... 10
    3.2.1 Calendar Dates ................................................................................................................... 10
    3.2.2 Times of Day .................................................................................................................... 13
      3.2.2.1 Time Zones ................................................................................................................ 15
    3.2.3 Units of Weeks .................................................................................................................. 16
    3.2.4 Duration ............................................................................................................................ 16
    3.2.5 Choosing Between Point and Duration Format .................................................................. 17
  3.3 Fuzzy Temporal Expressions ......................................................................................................... 18
    3.3.1 Past, Present and Future ..................................................................................................... 18
    3.3.2 Seasons ................................................................................................................................ 19
    3.3.3 Weekends .......................................................................................................................... 19
    3.3.4 Morning, Afternoon, and Night ....................................................................................... 19
  3.4 Modified Temporal Expressions ..................................................................................................... 21
    3.4.1 Language Specific Differences ........................................................................................... 24
      3.4.1.1 ON_OR_BEFORE and ON_OR_AFTER tokens ......................................................... 24
      3.4.1.2 Structural Position ....................................................................................................... 25
      3.4.1.3 Separated modifiers ..................................................................................................... 26
      3.4.1.4 The preposition « .............................................................. ........................................ 26

4 Determining the Extent of the Annotations .................................................................................. 28
  4.1 Lexical and Morphological Criteria ............................................................................................. 28
    4.1.1 Lexical Criteria ................................................................................................................... 28
    4.1.2 Morphological criteria ......................................................................................................... 29
      4.1.2.1 Pronouns ..................................................................................................................... 30
      4.1.2.2 Copula verb .................................................................................................................. 31
      4.1.2.3 Object and topic marker « ......................................................................................... 31
  4.2 Syntactic Criteria .......................................................................................................................... 32
    4.2.1 Appositives .......................................................................................................................... 35
    4.2.2 Range Expressions ............................................................................................................... 35
    4.2.3 Conjoined Expressions ......................................................................................................... 36
    4.2.4 Embedded Expressions ......................................................................................................... 37
      4.2.4.1 When to Create One Tag .............................................................................................. 37
      4.2.4.2 When to Create Multiple Tags, with Embedding ....................................................... 39
      4.2.4.3 When to Create Multiple Tags, without Embedding ................................................ 40
List of Tables

Table 2-1 Sample Lexical Triggers and Non-Triggers ................................................................. 2
Table 2-2 Sample Non-Triggers .................................................................................................. 4
Table 3-1 Temporal Expressions with the *Ezafe* ...................................................................... 9
Table 3-2 Present, Past, Future Tokens ..................................................................................... 18
Table 3-3 Modifier Tokens ........................................................................................................ 22
1 Introduction

This document is a supplement to the TIDES Standard for the Annotation of Temporal Expressions, and is designed to assist system developers and annotators working with Persian (Farsi) language data.

This supplement consists of three parts:

Chapter 2. Determining What Kind of Expressions to Annotate

This corresponds to Chapter 3 of the TIDES Standard for the Annotation of Temporal Expressions, providing Persian examples of markable and non-markable expressions.

Chapter 3. Capturing the Meaning of Temporal Expressions

This section corresponds to Chapter 4 of the TIDES Standard for the Annotation of Temporal Expressions, discussing language-specific issues. This section follows the order provided in the original document and presents specific examples either where Persian differs from English in expressing time or when it was felt that additional language-specific information could clarify the annotation guidelines. The cases that are similar to English, however, are typically not discussed and one should refer directly to the original document for how to normalize the temporal expressions using the attributes of the <TIME2> tag.

Chapter 4. Determining the Extent of Annotations

This section corresponds to Chapter 5 of the TIDES Standard for the Annotation of Temporal Expressions, providing Persian examples for establishing where the temporal expressions begin and end. Language-specific differences are also discussed.

---

1 The document includes numbers in both Persian and Latin numerals. In order to be able to see the Persian numbers correctly, please set the Numeral feature to ‘Context’ on MS Word. On MS Office 2007, this is achieved by selecting the Office button, choose Word Options, select Advanced, and under Show document content, set Numeral to ‘Context’ (Numeral is by default set to ‘Arabic’ referring to ‘European-Arabic’).
2 Determining What Kinds of Expressions to Annotate

This annotation standard is focused on temporal expressions. Such expressions can reference calendar dates, times of day (TOD), or durations (such as periods of hours, days, or even periods of centuries). Basically, if a phrase or word refers to some area on a timeline, we want to capture its meaning.

2.1 Markable Expressions

Markable expressions are the expressions that should be annotated. To be markable, the syntactic head of the expression must be an appropriate lexical trigger. Each lexical trigger is a word or numeric expression whose meaning conveys a temporal unit or concept, such as "month" or "daily".

<table>
<thead>
<tr>
<th>Part of Speech</th>
<th>Lexical Triggers</th>
<th>Non-Triggers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>دقیقه، ساعت، ثانیه، بند از ظهر، نصف شب، شب، روز، پیروزی، شبانه‌روز، بامداد، ماه، سال، تاریخ، قبل، جریان، جهان، سالن، دوره، موقع، مرکز، آینه</td>
<td>لحظه، دم، مناسبت، فرصت، موقعیت، زمینه، برنامه، مورد، دفع، ابزار، سال نوری، تاريخ (when meaning “history”)</td>
</tr>
<tr>
<td>Proper name</td>
<td>دوشنبه، فوریه، ابان، نوروز، عیدفطر، شب بلدا، کریسمس</td>
<td></td>
</tr>
<tr>
<td>Specialized time patterns</td>
<td>8:00:00، 1994، 30/12/2007، 2005</td>
<td>بعدی، قبلی، دیگر، معاصر، مدرن، هدیه‌ای، هم‌زمان، جدید، مکرر، بیوسته، زود، متعاقب، ابتدایی، مامتی، هم‌زمانی</td>
</tr>
<tr>
<td>Adjective</td>
<td>گنگشته، جاری، انتهای، اخیر، سابق، گذشته، آنی، روزانه، پیش‌تر، بعد، دیروزی، امروزی</td>
<td>فوراً، آن‌ها، هم‌زمان، پیش‌بینی، پیش‌بینی، بالاخره، بلافاصله، دوباره، قبل، بعد، احیاناً، معمولاً، مکرر، هم‌زمان، هنوز، زودی، مادر، عمر، سرانجام، عاقبت، بجا، بوانا</td>
</tr>
<tr>
<td>Adverb</td>
<td>فعلاً، اخیراً، سالانه، ماهانه، شوشته، امروزه، هر‌ساعت، کوتاه‌مدت، دراز مدت</td>
<td></td>
</tr>
<tr>
<td>Time noun/adverb</td>
<td>الآن، حالاً، دیروزی، امروزی، اینک، زمان‌ها، اکنون، هم‌زمانی</td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>2 (مثال: ۲ می‌رسیم)، پنجم (مثال: پنجم زانویه)، شصت (مثال: دهه شصت)</td>
<td></td>
</tr>
</tbody>
</table>
To be a trigger, the referent must also be able to be oriented on a timeline, or at least oriented with relation to a time (past, present, future). Table 2-1 contains a sampling of lexical triggers. The table also includes examples of closely related temporal concepts that are not considered triggers. Note that the same term can be a lexical trigger in some contexts and not in others since many words have ambiguous meanings. It is therefore vital for the annotator to consider the context of temporal expressions, rather than simply being triggered by the string. For instance, the term "today’s date is February 20th", but it is not a trigger in “the history of Islam and Iran” where it has no temporal meaning. These non-markable terms are discussed further in Section 2.2.

Markable expressions also include pronouns (such as آن) that can co-refer with a markable time expression. See the section “Pronouns and Elided Elements” in the TIDES Standard for the Annotation of Temporal Expressions.

2.2 Non-Markables

In contrast to the triggers presented above, the non-triggers (i.e., the non-markables), although they can be temporal in their semantics, are as a class less amenable to being pinned down to a timeline. For practical reasons, this ability to orient an expression in time is the basic distinction we rely on in constraining the scope of markable expressions. Of course, even the designated markables can be non-referring (and thus impossible to relate to a timeline), as in idiomatic phrases like "don’t count your chickens until they hatch" (lit. “one counts the chickens at the end of fall”) or “many happy returns” (lit. “one hundred years to these years”).

Likewise, note that adjective non-triggers are permitted within the extent of a markable expression, as in “in the modern era”, “the previous week/the week before that” or “in a few more days”. They are not markable on their own, as in نویسندهان “the modern writers”. To give another example, the adjective listed as a non-trigger in Table 2-1 would be included in the markable expression "the next day" but it will not be marked in "the next decision” where the syntactic head is “decision” and therefore not a lexical trigger for temporal annotation. Hence, and معاصر are not intrinsically terms that indicate a temporal expression, but whether they are markable or not depends on the context in which they appear (e.g., if the noun that they modify is a lexical trigger). However, starting with the 2005 guidelines the extent may also include non-triggers that are syntactic heads of markable expressions when the semantic head is a trigger word; e.g., partitives like بیشتر هفته “much of the week”.

2.2.1 Non-Markable Parts of Speech: Prepositions and Subordinating Conjunctions

Table 2-1 above shows that some parts of speech contain both triggers and non-triggers. In contrast, subordinating conjunctions which introduce clauses are never triggers; that is, they
never appear as the syntactic head of an annotated expression. The following table shows some examples of this part of speech. In addition, prepositions that introduce noun phrases are not considered triggers and are excluded from the annotated expression. These are exemplified in Table 2-2 as well.

### Table 2-2 Sample Non-Triggers

<table>
<thead>
<tr>
<th>Part of Speech</th>
<th>Non-Triggers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subordinating Conjunction</td>
<td>وقتی که، زمانی که، در حالیکه، همینکه، نامعلومی که، در ضمن اینکه، هر دفعه(که)، هر وقت که، زودتر از اینکه</td>
</tr>
<tr>
<td>Preposition</td>
<td>در، از، تا، طی، در خلال، حین، قبل از، بعد از، پس از، پیش از، در طول</td>
</tr>
</tbody>
</table>

Note that in Persian, conjunctions often contain که. In addition, in conversational Persian که can be used to mean “when” which is also a conjunction and should not be marked as in “when I was young…”.

In Persian, there are two types of prepositions: those that are derived from nouns and participate in the *ezafe* construction\(^2\) (e.g., طی این دو ساعت “during these two hours”), and a limited set of prepositions that do not take the *ezafe* (e.g., تا سال گذشته “until last year”). We treat both instances as a preposition and do not include them within the underlined annotated expressions. This can give rise to some ambiguities as most ‘nominal prepositions’ can also be used as simple nouns. Hence in the phrase در این مدت “in this period”, مدت is used as a temporal noun and is a lexical trigger. However, in the phrase به مدت دو سال “for two years” the term مدت is part of the compound preposition به مدت “for” and is not included in the temporal expression.

In certain instances, the preposition behaves more like a modifier of the temporal expression; in these cases, the preposition should be included within the annotated expression. See Section 3.4 “Modified Temporal Expressions” for examples.

#### 2.2.2 Non-Markable Point and Duration Expressions

Each of the expressions illustrated here only vaguely indicates a point in time (calendar date or time of day), or references some vague duration (interval) of time. Although many point and duration expressions are markable, the ones illustrated here are not.

**Sequencing and ordering** expressions are not markable:

- دولت پاکستان قبل گفتہ بود کہ انتخابات عمومی برگزار نخواهند شد.
  The Pakistani government had previously said that the general elections will not take place.

- معرفی برترین وبلاگ‌های فارسی همزمان با جشن تولد پرشن بلگ بود.
  The introduction of the best Persian weblogs was simultaneous with Persian Blog’s anniversary.

- سپرایانو ورودیس فارسی بعد از ماه‌ها تلاش تیمی در دسترس عموم قرار گرفت.
  After months of teamwork Persian Wordpress has finally been released to the public.

---

\(^2\) The *ezafe* links elements of the noun phrase and preposition phrase. It is discussed in more detail in Section 3.1.
Art will eventually reveal who we used to be and who we have now become.

This explosion attracted a large crowd and the subsequent explosion caused injuries to a number of them.

Those who have young children do not know what future awaits them (lit. what future they have ahead).

These expressions are admittedly borderline in terms of markability. With the anchoring attributes ANCHOR_VAL and ANCHOR_DIR, it might well be feasible to capture some of the semantics of these terms. For example, consider the following sentence:

Richard Haass said today in an interview with the Newsweek weekly that a difficult season lies ahead for Pakistan.

This sentence may be tagged as shown below (where the reference date is the 22nd of October, 2007). For more information on Anchoring, see Section 4.2.4 of the TIDES Standard for the Annotation of Temporal Expressions.

There are many ordering/sequencing terms that would need to be more fully explored before they could be added to the list of markables. Thus, although such terms are not “officially” markable, we encourage TIME2 users to explore this possibility and report their results to the NLP community.

Manner adverbs, which say how soon or how quickly something is done, are not markable:

I was just about to come look for you. / I was going to come look for you momentarily.

Non-quantifiable durations are not markable:

This key deletes the parameters permanently without placing them in the recycle bin.

3 Consider:
Negatives and references to non-existent times are not markable:

وزارت امورخارجه آمریکا اعلام کرد هنوز هیچ تاریخی برای مذاکره با ایران درباره اوضاع عراق تعیین نشده است. The U.S. State Department announced that no date has yet been set for talks with Iran regarding the situation in Iraq.

“Time” when it means “Situation” or “Occasion” is not markable. In Persian, the meaning of “situation” or “occasion” is often indicated by the words بار or مربیه دفعه. These words should not be marked:

به نظر من تابستان هم فرصت خوبی برای تفریح است و هم فرصت خوبی برای درس خواندن.

In my view, summers are both a good time for having fun and good time for studying.

بار اولین نیست؛ اما گفتم شاید این دفعه عوض شده باشه.

It’s not his first time; but I thought maybe he has changed this time.

Although لنحوه is generally translated as “second” or “instant”, it can also mean an “occasion” or “opportunity” as in the following sentence where it is not markable.⁴

We have not lost a giant opportunity as big opportunities are still and always ahead of us.

2.2.3 Non-Markable Frequency Expressions

Bare frequencies (frequency expressions with no time period given) are not markable:⁵

جرا انتقال گیبیهای مکرر قدر خون توسط فرد مبتلا به مرض قند حانز اهمیت است?

Why are frequent measurements of blood sugar by the person who has diabetes so important?

منظورم اینه که کم به خاطر شروع ترم جديد دیگه نمیتونه زیرنشود آپ کنم.

What I mean is that gradually because of the start of the new term I won’t be able to upload frequently anymore.

بارها گفت‌ه‌به نبودن که قهرمانی روش‌نوسنده نیست.

I had said time and time again that heroism is not the style of the writer.

منطقه‌ی غرب معمولب‌ا بی‌محافظه‌کاران رای‌می دهد.

The West usually votes for the conservatives.

همیشه‌ه راه‌نarnings هست.

There is always another way.

Other examples in this class include:

به‌نفس‌ه، معمولب‌ا، بطور کلی، بطور عادی، بطور معمول، معمولب‌ا، بطور‌دمام‌، بطور‌رد، مکرر، دانمار، هزغ‌ز، گانی، گان‌گاها، اغلب.

2.2.4 Non-Markable Proper Names

Proper names that designate something other than a temporal entity but happen to contain lexical triggers are not markable. The following list contains names of organizations, books, festivals,

⁴ Note that in this example هنوز (still) and همیشه (always) are also not markable.

⁵ In contrast, frequencies whose semantics include a temporal unit, such as هر روز (daily) or زمانه (at a time), are markable. See section 4.5 in the TIDES Standard for the Annotation of Temporal Expressions.
albums, and movies (shown underlined). Since these are not temporal entities, they are not markable. Some of these expressions may be ambiguous in Persian since proper names are not capitalized and punctuation such as quotation marks is not generally used; annotators should use the context of the sentence to determine the correct marking.

The Palestinian organization *Black September* claimed responsibility.

Doesn’t anyone have George Orwell’s work *Nineteen Eighty Four* in Persian?

Kiarostami did not participate in the *Fajr* Film Festival.

Third Day is Mohamad Hossein Latifi’s latest directorial film.

Linkin Park is supposed to issue its third album *Minutes to Midnight* on 15 April 2007.

Let us believe in *the beginning of the cold season* is a collection of poetry from Forugh Farrokhzad and includes 7 poems.

However, triggers that are functioning as temporal modifiers within titles (as opposed to proper names) are markable. Examples include titles of conferences and awards:

The 2002 Pulitzer Prize (lit: The Pulitzer prize of year 2002)

6 Note that *fajr* is a temporal expression according to the *TIME2 Persian Calendar Extensions* as it represents the time of the morning prayer between the break of dawn until sunrise.
3 Capturing the Meaning of Temporal Expressions

This section is a supplement to Section 4 in the TIDES Standard for the Annotation of Temporal Expressions which illustrates how the semantics of markable expressions are captured in the annotations. The focus of this section is to provide Persian language-specific examples that may raise issues for annotators.

3.1 General Notes on Persian

3.1.1 Ambiguities in Persian

As noted in Section 2.2, many expressions may be markable or not depending on context (e.g., does it refer to a particular time, or is it a non-temporally-anchored reference to an occasion?). It is therefore important to always consider the context of temporal expressions in a text.

For instance, the term حضور الآن may mean momentarily as we have seen in Section 2.2.2, in which case it should not be annotated.

I was just about to come look for you. / I was going to come look for you momentarily

However, it can also mean right now or nowadays as shown below. In these cases, حضور الآن is a temporal expression and should be tagged.

What are you listening to right now?

Nowadays, the political prisoner is under various types of psychological pressures and tortures in Iran.

3.1.2 Syntax of Noun Phrases

An important distinction in Persian, as compared to English, is the use of the اضافه (ezafe) in the formation of the noun phrase. The noun and its modifiers (i.e., adjectives) and possessors (i.e., possessive pronoun or possessor nouns) are linked to each other by using an affix known as the ezafe, which is pronounced as /e/ after consonants and /ye/ after vowels. In English, the adjectives simply precede the noun and sometimes ‘of’ is used to express possession, whereas in Persian the ezafe is used in both cases, as shown in the contrasting examples below:

“small black sparrow”

“Nima’s small black sparrow / the small black sparrow of Nima”
The ezafe construction plays an important role in the formation of temporal expressions as can be seen in the following parallel Persian and English examples. Each noun phrase in this table is a temporal expression and should be annotated as a unit.

**Table 3-1 Temporal Expressions with the Ezafe**

<table>
<thead>
<tr>
<th>Persian Expression and Gloss</th>
<th>English Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>عصر روز چهارشنبه</td>
<td></td>
</tr>
<tr>
<td>evening-Ez day-Ez wednesday</td>
<td>Wednesday evening</td>
</tr>
<tr>
<td>روزهای اینده</td>
<td></td>
</tr>
<tr>
<td>days-Ez future</td>
<td>the future days/the upcoming days</td>
</tr>
<tr>
<td>سال 1342 خورشیدی</td>
<td></td>
</tr>
<tr>
<td>year-Ez 1342-Ez solar</td>
<td>1342 AP</td>
</tr>
<tr>
<td>فصل زمستان امسال</td>
<td></td>
</tr>
</tbody>
</table>
| season-Ez winter-Ez this.year | this year’s winter  
|                                      | = the winter of this year |
| سال های ابتدای انقلاب  |
| years-Ez early-Ez revolution  | the early years of the revolution |
| ساعت 8 شب دوشا بهته چهارم این ماه  |
| hour-Ez 8-Ez night-Ez monday-Ez week-Ez fourth-Ez this month | 8 o’clock on monday night of the fourth week of this month |

The difference in the noun phrase constructions in these two languages is crucial especially with respect to Section 4.2.4.1 “When to Create One Tag” (cf. Section 5.2.4.1 of the TIDES Standard for the Annotation of Temporal Expressions) as many instances that use a preposition in English are actually linked via the ezafe in Persian and therefore do not raise similar issues:

1. The preposition “of” in temporal expressions in English often corresponds to the ‘ezafe’ in Persian:

   دوم سپتامبر
   the second of December

   تابستان 1964
   the summer of 1964

   صبح یازدهم سپتامبر
   the morning of September 11th

2. Partitives formed in English using the “of” preposition are typically formed with the ‘ezafe’ in Persian:

   تمام سال
   all of the year
3. The preposition “in” in TOD expressions in English corresponds to the ‘ezafe’ in Persian:

پایده صبح
eleven in the morning

4. Prepositions used with “early” and “late” expressions in English correspond to the ‘ezafe’ in Persian:

اوائل آن سال
earlier in the year (cf. “earlier that year”)

ظهرت شب
late at night

All of these noun phrase constructions should receive a single TIME2 tag when annotating the temporal expressions in Persian.

3.1.3 Conversational Persian

There is an instance of diglossia in Persian where both a literary dialect and a spoken, conversational variant coexist. Both may be used in writing depending on the author, topic, and document style or register. Hence the conversational or colloquial variant will be found more commonly in blogs, chats, dialogue, and personal correspondence. The examples provided here are extracted from both variants of the language. Furthermore, if an expression is markable in literary Persian, then its conversational variant and equivalent is also markable, and vice versa.

3.2 Precise Temporal Expressions

3.2.1 Calendar Dates

Persian language documents may use several alternative calendar systems, sometimes within the same text. The attribute ‘CAL’ should be used in the TIME2 tag to identify the calendar system being used. For detailed information, see the TIME2 Persian Calendar Extensions and TIME2 Islamic Calendar Extensions. The Persian calendar is a solar calendar and is used as the official calendar system in Iran and Afghanistan. In Persian, Anno Persicum or AP is indicated alternatively as ‘hejri-e khorshidi’ (هجری خورشیدی), ‘hejri-e shamsi’ (هجری شمسی), ‘khorshidi’ (خورشیدی) or ‘shamsi’ (شمسی). In Persian, the Gregorian calendar or the Christian year (CE) is
called ‘Miladi’ (ميلادي). If the Islamic lunar calendar or AH is used in a document, it is referred to as ‘hejri-e ghamari’ (هجری قمری or simply as ‘ghamari’ (قمری).\footnote{The common abbreviations used are م for ميلادي, ه for هش, ق for قمری, and ه for هجري.}\footnote{7}

The type of calendar system being referred to may be explicitly stated in the text as in the following examples (assume for all these examples that the reference date is Monday, October 22, 2007 which is equivalent to 30 Mehr 1386 AP):

\begin{verbatim}
roz dham sitamir sall jari milad
The tenth of September of the current Christian year
<TIME2 CAL="ISO_EXT" VAL="2007-09-10"> roz dham sitamir sall jari milad
</TIME2>
qiym behem 1357 hejri xoroshidi
The rebellion of Bahman 1357 AP
<TIME2 CAL="PERSIAN_FA" VAL="1357-11"> qiym behem 1357 hejri xoroshidi</TIME2>
derasal 1372 shamsi
In 1372 AP
<TIME2 CAL="PERSIAN_FA" VAL="1372"> der sal 1372 shamsi</TIME2>
dr 1372 shamsi
The first of Moharram of 1429 AH
<TIME2 CAL="ISLAMIC" VAL="1429-01-01"> dr 1372 shamsi</TIME2>
heft thor 1358
7 Sawr 1358
<TIME2 CAL="PERSIAN_AR" VAL="1358-02-07"> heft thor 1358</TIME2>
penjashnine bist wa som anzar 1385
Thursday twenty third of Azar 1385
<TIME2 CAL="PERSIAN_FA" VAL="1385-09-23"> penjashnine bist wa som anzar 1385</TIME2>
akhtar 1492
8 October 1492
<TIME2 CAL="ISO_EXT" VAL="1492-10-08"> akhtar 1492</TIME2>
awlolin rooz mahi
\end{verbatim}
The first day of the month of May

Some documents may list alternative calendar dates referring to the same day. In these cases, each instance should be tagged separately as shown:

- Wednesday 29 Farvardin 1386 AP corresponding to 18 April 2007 CE equivalent to 29 Rabi-ol-Awwal 1428 AH

If no calendar system has been mentioned in the context of the document, annotate the text according to the default (i.e., Gregorian) system.

3.2.1.1 Decades, Centuries, Millennia, and BCE

As with regular dates, the calendar system used to represent decades (قرن، سده،دهه)، centuries (هلاله) and millennia (هزاره) are generally explicitly marked or can be determined by the context of the document.

Similarly, the pre-calendar dates may be explicitly tagged for the calendar system used as in

ٖدؽی نّكی نةعٙت پؽٚؼ ثٛظٖ اقت.

An environmental crisis in the third millennium!
The art of carpet weaving in Iran goes back to 5000 years ago.

To simplify the annotations in the rest of this document, the CAL attribute will not be included in the following examples unless the Persian or Islamic calendar is explicitly used.

### 3.2.2 Times of Day

In formal contexts, the 24-hour time period may be used to express time in Persian. However, “a.m.” and “p.m.” are most often expressed using words for the part of day such as صجر, قبیظل١م for “morning”, ظصؽٙؽ for “afternoon”, عص for “evening”, or ظث for “night” (for all examples, assume that the reference date is October 22, 2007).

#### جهارمين نشست

The fourth session took place at 15:00 on Thursday, the 22 of Azar in the Amphitheatre Hall of the Center.

#### صجر 8 ٪ عص

It’s 8:30 in the morning and I have had two cups of tea and have discussed unemployment with couple of people.

#### ظصؽ ظغ کرگچ

On Friday the water of Tehran will be interrupted in all areas from 5 to 7 in the evening.

#### ظث 9 ٪ عص

The city’s trash collection takes place at exactly 9 o’clock at night.

In Persian, the relative placement of the word “night” next to the name of a day is extremely important as it may refer to two distinct days. Hence, جمعهشب (lit. Friday night) refers to “Friday night” while شبجمعه (lit. night of Friday) means “Thursday night”.

---

9 In this sentence, راس (exactly) is the modifier of the temporal expression and is therefore included in the annotation. See Section 3.4 “Modified Temporal Expressions” for discussion.
The words نیمهشب and to a lesser extent نصف شب are used to express “midnight” in Persian, but they actually translate into “middle of the night” and are therefore ambiguous. The annotator needs to use the context to determine the time meant by the author. Hence, midnight can be expressed as follows:

عزاداري محرم، بعد از ساعت 12 شب ممنوع شد.

The mourning for Muharram is prohibited after midnight (lit. after 12 at night).

دوباره زن وشوهر کنار هم نشستند و ساعت تیلداک میکرد؛ باالاخره نیمه شب شد.

Husband and wife were once again sitting next to each other and the clock was ticking; it finally turned midnight (lit. middle of the night)

Nevertheless, both نیمهشب and نصف شب are often used to mean some time in the night (usually between 1am and 4am) and it can therefore cover a wide range of time expressions. In most instances, a specific time is used along with these expressions as shown in the examples below:

1:30 بعد از نیمه شب شد.
It turned 1:30 in the night.

3:00 بعد از نیمه شب.

We received this message at three in the night.

آینا ساعت سه وربع نصف شب است و هنوز خواب نبرده.

It is now three and a quarter in the night here and I am still not asleep.

The words، نیمه‌دوز، “midday” and ظهر “noon” can both be annotated as 12 o’clock:

بردیل به ساعت نگاهی انداد: درست نیمه‌دوز بود.
He glanced at his watch for no reason: it was exactly midday.

و نیمه‌دوز به ساعت نگاهی انداد: درست بود.

The term شبانه‌روز in Persian literally means “nightly day” and refers to the 24-hour period that forms a whole day; it refers to both daytime and nighttime. In the following example, شبانه‌روز (although referring to a 24-hour period as can be seen in the English translation) can be expressed as a specific day for the purposes of the TIME2 annotation.

شبانه‌روز

Eight people were killed in the unrest in the Armenian capital the past 24 hours.

See Section 3.3.4 on annotating usages of نیمهشب and نصف شب that do not seem to refer to a specific time.
3.2.2.1 Time Zones

To express different time zones in Persian, the explicit location is mentioned in the text. Therefore, the relative time of each location with respect to Universal coordinated Time (UTC) should be determined and indicated in the annotation. Hence, a “Z” in the first example indicates that the time is given in UTC or Greenwich Meridian Time, and the “+01” in the second sentence indicates that the local time (in this case, Central European Time) is one hour ahead of UTC.

A large aftershock with a magnitude of 5.3 occurred at 3 o’clock and 6 minutes and 32 seconds in the morning at Greenwich time.

The official start of the meeting is at 7 pm Central European Time.

The time for the transition to the new year 1386 AP will be at 9:18 on Thursday, Farvardin 1, 1387 at the time of Iran.

The funeral ceremony of the late Mahasti, the deceased Iranian artist, is taking place on Friday 29 June at 8 pm at the time of Stockholm (11 at the time of Los Angeles) at Westwood Memorial Park cemetery in Los Angeles.

Context may need to be used to determine the local time for annotation purposes as in the following example referring to voting booths in Oman on October 27th, 2007:

Note that Iran generally operates Daylight Savings Time (DST) between 1 Farvardin (March 21) and 1 Mehr (September 23) when the time is 4.5 hours ahead of Greenwich Mean Time (i.e., GMT+4:30), and the rest of the year it’s 3.5 hours ahead of Greenwich time (i.e., GMT+3:30). However, similar to the annotation of time zones in Section 4.2.2.1 of the TIDES Standard for
the Annotation of Temporal Expressions, if DST is not overtly indicated in the expression, it does not need to be included in the tag.

3.2.3 Units of Weeks

If a Persian or Islamic calendar is used in the document and the text refers to a week-based temporal expression, follow the guidelines described in the TIME2 Persian Calendar Extensions and TIME2 Islamic Calendar Extensions. As an example consider the temporal expression below:

This article was written on the Monday of the fourth week of Ordibehesht of 1386.

As Ordibehesht is the second month of the year in the Persian calendar system, week 4 of Ordibehesht would coincide with week 8. Monday is day 3 since day 1 of each week is Saturday.

3.2.4 Duration

An expression of duration indicates a period of time, indicating how long something lasts. Durations that refer to specific periods of time can be oriented or anchored with respect to other points or periods of time. The annotation of duration events in Persian is similar to the guidelines described for English, but three specific points can be emphasized:

The term نمبشوز in Persian refers to the 24-hour period that forms a whole day; it refers to both daytime and nighttime. When referring to a duration، نمبشوز can be tagged as the TIME2 ‘Day’ token، as exemplified below where we translate نمبشوز as “a 24-hour period”.

A specialist of brain and nerve diseases said that on average each person should sleep seven and half hours during a 24-hour period.

WITHIN, STARTING and ENDING. As in English، an expression of age is markable when it is an adjective phrase. In Persian the expression would be formed with سال، which is itself formed on the noun سال “year” as in the examples below.

11 Note that week 01 of the year is the first week of Farvardin that contains a Tuesday according to the TIME2 Persian Calendar Extensions and given the fact that the year 1386 AP began on a Wednesday (i.e.، Farvardin 1 fell on a Wednesday)، the first week begins on Saturday، 4th of Farvardin and the fourth week of Farvardin begins on Saturday the 25th.
The 89-year-old poet passed away on Thursday night.

Choosing between STARTING and ENDING. The tense used in the sentence can help determine whether an ongoing duration is involved. For instance, in the example below the use of the present tense [lit. “three day is”] instead of the past tense [lit. “three day was”] suggests that the strike is still continuing and the annotator should count back three days to enter the ANCHOR_VAL.

If the sentence indicates both a start and an end date, the STARTING value is preferred:

3.2.5 Choosing Between Point and Duration Format

Whether something is considered a duration or a point in time can depend largely on the context. Almost identical expressions can be tagged differently if the context implies different meanings. The following two examples in Persian show the same word used in two different contexts:

Point in Time:

Duration:

In the first example above, بعد means “later” and modifies the temporal expression; in the second example, it is part of the compound preposition بعد از “after” and is not included in the annotation.
of the temporal expression. Thus, it is often the context, and not the expression itself, which dictates the format of the annotation.

3.3 Fuzzy Temporal Expressions

3.3.1 Past, Present and Future

Many temporal expressions refer in general terms to the past, the present, or the future. These can be expressed with the tokens listed in the table below, which also includes sample markable and non-markable expressions.

<table>
<thead>
<tr>
<th>Token</th>
<th>Sample Markable Expressions</th>
<th>Sample Non-Markable Expressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRESENT_REF</td>
<td>(with the “nowadays” meaning) آینه، اکنون، کنونی، امروزی (the “nowadays” meaning) انّ، ہمینّالآنّ، حالّاً</td>
<td>فوراً، آناً</td>
</tr>
<tr>
<td>FUTURE_REF</td>
<td>آینه، روزهای آینده، چند سال/ماه/روز دیگر (meaning “in the future”) آینه، پیشگاه</td>
<td>ئیندا، پیشگاه</td>
</tr>
<tr>
<td>PAST_REF</td>
<td>گذشته، آخر، سابق، دیروز، این اواخر، ماهها/سالها پیش (at one time”, “once”) زمانی، یہ موقعی</td>
<td>قبلًا، قبل از این، سابقاً</td>
</tr>
</tbody>
</table>

Several examples are shown below, but in general the annotation guidelines for these tokens follow the guidelines described for English in Section 4.3.2 of the TIDES Standard for the Annotation of Temporal Expressions.

مهاجرین دیروز، از مهاجرین امروز حمايت کنید! کنید! Yesterday’s immigrants, support today’s immigrants!

از مهاجرین دیروز

زمانی، یہ موقعی

<TIME2 VAL="PAST_REF" ANCHOR_DIR="BEFORE" ANCHOR_VAL="2007-10-22"></TIME2>

<TIME2 VAL="PRESENT_REF" ANCHOR_DIR="AS_OF" ANCHOR_VAL="2007-10-22"></TIME2>

The sentence below, repeated from Section 3.1.1, exemplifies the use of ہمین الآنّ to mean “nowadays” or “these days”:

زندانی سیاسی در ایران در زیرانواع فشارها وشکنجه های جسمی وروحیست.

<TIME2 VAL="PRESENT_REF" ANCHOR_DIR="AS_OF" ANCHOR_VAL="2007-10-22"></TIME2>

Nowadays, the political prisoner is under various types of psychological pressures and tortures in Iran.
The following sentence discusses an interview with Mr. Modarressi carried out on 14 December 2006, and the author claims that Mr. Modarressi has countered his own position just a few days earlier in a written article. Thus 14 December 2006 can be used as the Anchor value in this example:

Mr. Modarressi himself, in a previous article he has written a few days earlier on this topic, has taken a 180 degree stance against his 14 December interview.

3.3.2 Seasons

Note that for the Persian calendar, the seasons align perfectly with the year so that no season falls in between two years (as does Winter in the Gregorian calendar).

3.3.3 Weekends

For a description of how to annotate weekends in a Persian calendar system, see the TIME2 Persian Calendar Extensions.

3.3.4 Morning, Afternoon, and Night

Periods of the day are often used vaguely in Persian. For instance, نیمیشنب and نیم‌نشاب can be used to mean “midnight” or they can refer to a time at night typically between 1 and 5 a.m. (also see examples in Section 3.2.2). Note that in the sentences in this section, only the English translations of نیم‌نشاب and نیمیشنب are shown in italic, and other temporal expressions are ignored. In the two examples below, نیم‌نشاب and نیمیشنب refer to “midnight”:

I suddenly looked at the watch and saw that it was past midnight.

I suddenly looked at the watch and saw that it was past midnight.

The code for Operation Ajax was supposed to be announced as a code from Radio BBC on midnight of that day under the name of “the time is exactly midnight”.

However, in general نصف شب and نصف صبح mean any time between midnight and sunrise, so they are best translated into English as “a.m.”:

آی اس پی من نرخش اینه: ساعت 6 تا 1 نیمهشب 100 توم؛ ساعت 1 نیمهشب تا 8 صبح 90 توم.
The rate for my ISP is this: 6 to 1 am 100 tumans; 1 am to 8 am 90 tumans.

[Lit.: The rate for my ISP is this: 6 to 1 middle of the night 100 tumans; 1 middle of the night to 8 of morning 90 tumans]

And in some rare cases, نصف شب and نصف صبح may even be interpreted as starting after dusk when the dark sets as in the examples below. Note that in this case the tokens would span a new day boundary.

These terms can therefore be tricky where, depending on the context, they can be used as indicating midnight exactly and at other times, they refer to a range in the night. The tagged examples below show the annotation for each instance.

(1) نصف شب and meaning midnight

یه موقع به ساعت نگاه کردم دیدم از نصف شب گشته.
I suddenly looked at the watch and saw that it was past midnight.

یه موقع به ساعت نگاه کردم دیدم از نصف شب گشته. <TIME2 VAL="T24:00">نصف شب</TIME2> گشته.

قرار بود در نیمه شب آن روز رمز عملیات آراکس با عنوان «حالا دقیقا نیمه شب است» به صورت رمز از رادیو بی‌بی‌سی اعلام شود.
The code for Operation Ajax was supposed to be announced as a code from Radio BBC on midnight of that day under the name of “now the time is exactly midnight”.

قرار بود در نیمه شب آن روز <TIME2 VAL="1953-08-18T24:00">ركز عملیات آراکس با عنوان</TIME2>  «حالا دقیقا نیمه شب است» به صورت رمز از رادیو بی‌بی‌سی اعلام شود. <TIME2 VAL="1953-08-18T24:00">حالا دقیقا نیمه شب</TIME2> اعلام شد.

(2) نصف شب and corresponding to the English “a.m.”

آی اس پی من نرخش اینه: ساعت 6 تا 1 نیمهشب 100 توم؛ ساعت 1 نیمهشب تا 8 صبح 90 توم.
The rate for my ISP is this: 6 to 1 am 100 tumans; 1 am to 8 am 90 tumans.

As the time for the evening prayer actually begins after sunset and lasts until sunrise, “middle of the night” here can refer to a time before midnight, although the general native speaker judgment is to treat these words as referring to a time after midnight.
and نصف شب (3) meaning some time in the middle of the night (between midnight and sunrise). Although these terms do not directly correspond to the concept of “night” in English as they generally involve a time after midnight only, we will tag them as TNI (see Section 4.3.7 of the TIDES Standard for the Annotation of Temporal Expressions for the annotation of periods of the day).

Yesterday night in the foggy streets, we were searching for him. [Lit. In the middle of the night of yesterday …]

In the following example, the author uses the same term, نصف شب, to refer to both midnight and middle of the night (referring to a time between midnight and sunrise), clearly showing that the annotator needs to take context into account when tagging these items.

3.4 Modified Temporal Expressions

If a temporal expression is quantified or modified in some way, the modifying element should be expressed in the annotation. For example, 1386 سال “1386” is an unmodified expression, but late 1386 is modified. In general, we want the annotation to capture the basic semantics of quantifier modifiers (e.g., حدود، تقريبًا) and lexicalized aspect markers (e.g., اوایل، اغاز). Table 3-3 represents some of the modifier tokens in Persian.

We do not want to capture the semantics of leading prepositions or other terms that are outside the extent of the tagged temporal expression (see Section 4 of this document on Determining the Extent of the Annotations). For example, the expression بعد از دو ساعت “after two hours” is not considered a modified expression for our purposes because the compound preposition بعد از “after” is not included within the extent of the TIME2 tag. However, the expression بعد از دو ساعت بیش از دو ساعت “more than two hours” is a modified temporal expression and the adjective + preposition unit بعد از دو ساعت “more than” is included within the extent of the tag. In this case, the MOD attribute is used to capture the semantics of the modifier within the scope of the TIME2 expression. The distinct annotations for these two examples can be seen below:
Table 3-3 Modifier Tokens

<table>
<thead>
<tr>
<th>Points and Durations</th>
<th>Token</th>
<th>Sample Expressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points</td>
<td>BEFORE</td>
<td>بيش از (بيش از یک کسد سال قبل)</td>
</tr>
<tr>
<td></td>
<td>AFTER</td>
<td>کمتر از (کمتر از یک دهه بيش)</td>
</tr>
<tr>
<td>Durations</td>
<td>LESS_THAN</td>
<td>کمتر از (کمتر از 14 ساعت)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>قريب (قریب 28 سال)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>نزدیک به (نزدیک به هده سال)</td>
</tr>
<tr>
<td></td>
<td>MORE_THAN</td>
<td>بیشتر از (بیشتر از هشت ماه)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>بالاتر از (بالاتر از 18 سال)</td>
</tr>
<tr>
<td></td>
<td>EQUAL_OR_LESS_THAN</td>
<td>حداقل (حداقل 15 روز کاری)</td>
</tr>
<tr>
<td></td>
<td>EQUAL_OR_MORE_THAN</td>
<td>حداقل (حداقل یک دقیقه) اقل (اقلا دول)</td>
</tr>
<tr>
<td>Points and Durations</td>
<td>START</td>
<td>اغاز (اغاز سده بیست و یکم)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>شروع (شرو ست سال 1387)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ابتدای (ابتدای قرن بیستم)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>اولین (اوائل زمستان همان سال)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>سپهند (سپهند سال دوم میلادی)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>زود (صبح زود)</td>
</tr>
<tr>
<td></td>
<td>MID</td>
<td>اواست (اواست این ماه)</td>
</tr>
<tr>
<td></td>
<td>END</td>
<td>پایان (پایان قرن هشتم ق.م)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>آخر (آخر تابستان امسال)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>اواخر (اوخار این 84)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>دوروقت (دوروقت شب)</td>
</tr>
<tr>
<td></td>
<td>APPROX</td>
<td>حدود (حدود 12 سال بیش)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>تقییا (تقییا ساعت 9 صحیح)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>کم و بیش (کم و بیش 14 سال)</td>
</tr>
</tbody>
</table>

The following are several sentences exemplifying the usage of the various modifier elements with temporal expressions.

"A bit more than a hundred years ago, Britain, Germany and Italy sent a joint fleet to the coast of Venezuela.

13 See section 4.4.1 of the TIDES Standard for anchoring approximate offsets."
Human Rights Watch also wants the Iranian government to quickly and completely nullify the death penalty for all individuals who at the time the crime was committed were less than eighteen years old.

The board of directors is required to produce the company’s balance-sheet and profit and loss account and its revenue report to the auditor at least two months prior to the formation of the annual public assembly.

Iran today finds itself in the beginning of the twenty-first century.

Early in the morning, before eating anything, the level of glycogen in the muscles and liver is low.

The Iranian government arrested them approximately two years ago.

In addition, elements meaning only and exactly are considered modifiers and should be included in the extent of the annotation. However, they do not need to be represented with a MOD attribute (there are other temporal expressions in the following examples but only the expressions under discussion are highlighted):
On ChaharShanbe Suri exactly I finished my exams.

In the evening, right at six o’clock, I was in front of the Eastern gate of the Laleh Park.

Note that a number of the sample expressions in Table 3-3 can also be used as adverbs. These should not be included in the annotation, even if they appear next to the temporal expression, as shown in the examples below with “اللاً” “at least” and “ظ٠هت” “late” (shown in bold) as they are not modifying a temporal expression. Thus, the first sentence does not mean “I was happy at least today but also possibly in more future days” but rather that “I have generally not been very happy but at least, I was happy today”. Hence, “اللاً” is not a modifier of “today” but is a sentential adverb.

At least today I was happy...

Last night Hitchcock’s “The Birds” had amused me until late.

Haji, because he had too much to do, had to come home late on most nights.

3.4.1 Language Specific Differences

There are some issues specific to Persian that should be considered, mostly having to do with the distinct structures of Persian and English:

3.4.1.1 ON_OR_BEFORE and ON_OR_AFTER tokens

These tokens do not exist in Persian as there are no expressions such as “no less than” or “no more than” in this language. To express a similar concept, the token “less than” or “more than” has to be used along with a negation on the verb:

The validity of the prescriptions for syringes or drug solutions will not be for more than two days.

In these cases, there is no way of capturing the negation meaning in the annotation as it is not overtly stated in the temporal expression itself. Only the temporal expression should be annotated and tagged according to its appropriate MOD value.14

14 Note that constructions such as in the following example are not equivalent to the English sense of “no more than”. As the translation of نه shows, the negation here is the adverb meaning “not”; it is not part of the temporal expression.

I think it was for this very reason that Hossein Panahi has said I love life but not more than 20 years.
3.4.1.2 Structural Position

Some modifier tokens can appear either before or after the temporal element. For example, “late” can be following the noun as in the first two sentences below where it is linked to the head noun by the ‘ezafé’. It can also occur preceding the temporal noun as in the last example below; here, the two elements are not linked with the ‘ezafé’ yet “late” still modifies the temporal expression.

In the Tehran province from *late Monday night* which is tomorrow night the weather will be rainy and there will be light snow.

میکنید

“Sobhar” which is a new coinage is equivalent to the English “brunch” meaning a meal that you eat *late morning* and instead of breakfast and lunch.

Similarly, compound tokens that consist of the combination of an adjective and a preposition such as “less than” or “more than” can sometimes appear in different orders around the temporal expression. For instance, the most common way to say “more than ten years” would be [Lit. more than ten year] However, one may also encounter the construction [Lit. from ten year more], where the “more” and “than/from” parts have been separated but still appear near the temporal expression. These two constructions should be annotated differently: in the first sentence below, the modifier token is included in the extent of the annotation; in the second example, the preposition “from” is excluded from the annotation but the modifying adjective “more” is included. However, both constructions should include a MOD attribute defined as MORE_THAN.

More than twenty percent of this country’s workforce was unemployed for *more than the past ten years*.

میکنید

What I mean is that when you go to court and they give you *more than ten years* [Lit. from ten year more], I am at your service.
3.4.1.3 Separated modifiers

In some instances, the modifying token may not be adjacent to the temporal expression as in the following examples where the modifier is in bold and the relevant temporal expression is underlined (italic in the English translation):

niej سال خورشیدی تقیبیا معادل ۳۶۵ روز، ۵ ساعت و ۴۹ دقیقه است.
A solar year is approximately equal to 365 days, 5 hours and 49 minutes.

ابن تقویم دیواری دقیقا در روز اول سپتامبر ۲۰۰۸ بعنی ۱۱ شهریور سال ۱۳۸۷ منتشر خواهند شد.
This wall calendar will be printed exactly on the first day of September 2008 which is 11 Shahrivar of 1387.

جان فلیکس انتونی سندا در ۲۳ اوریل ۱۹۷۷ دیده به جهان کشیده تقدیبیا میشه سه اردیبهشت روز تولدش.
John Felix Anthony Cena was born on 23 April 1977, it’s almost on three Ordibehesht his birthday.

سه بار کمیسیون سنگین وزن که اقلا از اول سال دهه هفت دستش دود.
Three time heavyweight championship that he has held at least from the beginning of two thousand five until towards the end of two thousand eight.

In these cases, as in the similar instances in English, the modifier token should not be included in the extent of the annotation and the MOD value will not be represented in the tag.

3.4.1.4 The preposition 

The preposition تا can sometimes modify a duration expression when it is used in the meaning of “up to” in English. For instance, in the following sentence, تا is used to refer to an amount that is equal to or less than the temporal amount specified in the expression (i.e., in three years or less, the number of internet users will be 35 million people). In these cases, although prepositions are generally not included in the temporal expression tags, the preposition تا is treated as a modifier token. It is included within the extent of the annotation and should be represented with the MOD value “EQUAL_OR_LESS”.

هم اکنون تعداد استفاده کنندهان از اینترنت در ایران هفت و نیم میلیون نفر است ولی انتظار می رود تا سال دیگر این تعداد به سی و نیم میلیون برسد.
Currently, the number of Internet users in Iran is seven and half million people but it is expected that up to three years from now this number will read thirty five million.

The following are additional examples for the modifier usage of the preposition تا in Persian with annotations. Note that these always occur with duration expressions.

شخصی که از قبل دارای گواهینامه و یا تجربه رانندگی نباید تا یکسال حق دادن امتحان رانندگی را ندارد.
An individual who has not previously held a driver’s license or does not have driving experience does not have the right to take the driving test for up to a year.

این فرد که در این دوره گواهینامه و یا تجربه رانندگی داشته تا یکسال حق دادن امتحان رانندگی را ندارد.
He knew that his breadwinner will be in jail for up to six more months.

او می دانست که نان اورش تا شش ماه دیگر در زندان خواهد ماند.
And he knew that his breadwinner would be in prison until six more months.

که نان اورش تا شش ماه دیگر در زندان خواهد ماند.
He knew that his breadwinner would be in jail for up to six more months.

Perfectly.
This highly watched show was so constraining for those involved with it that for up to three years it did not allow them to participate in another picture.

In the following example, however, تا is not used as a modifier since the temporal expression is interpreted as until the year 2005 (but not including it).

The United Nations: The population of the world will reach 9 Billion by the year 2050.

The sentences below show further usages of the non-modifying usage of the preposition where تا usually means “until” or “up until” in English. In these cases, it is excluded from the extent of the annotation.\(^{15}\)

Note that another usage of the preposition تا is to present a range as in the sentences below, where it should also be excluded from the annotation (see Section 4.2.2 for more on Range Expressions).

\(^{15}\)There are other temporal expressions in these sentences; however, only the ones introduced by تا are underlined.
4 Determining the Extent of the Annotations

This section contains rules for determining where each temporal expression begins and ends. We refer to this as the *extent* of the expression.

4.1 Lexical and Morphological Criteria

4.1.1 Lexical Criteria

When tagged, the full extent of the tag must be one of the following grammatical categories in Persian:

- noun (including proper noun): e.g., امروز “today”, جمعه “Friday”
- noun phrase (NP): e.g., اون روزا “those days”, روز جمعه “Friday” (lit. “day of Friday”), اين دو سال آخر “the last two years”, يک ساعت تمام “a whole hour”, دو هفته پيش “two weeks ago”
- adjective: e.g., كنونی “current”
- adverb: e.g., اخيرا “recently”
- adjective/adverb phrase: e.g., تمکنجب ٔ١ُ قبٌٗ “nearly half hour”, پٕدبٖ ٚ ظٚ قبٌٗ “fifty-two year old”

The full extent cannot be a preposition phrase (i.e., the expression cannot start with a preposition) or a clause of any type (i.e., the expression cannot start with a subordinating conjunction). Thus, from the following phrases we get the annotated expressions listed below:

<table>
<thead>
<tr>
<th>Context</th>
<th>Extent of TIME2 Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>قبل از پنجشنبه “before Thursday”</td>
<td>پنجشنبه “Thursday”</td>
</tr>
<tr>
<td>از صبح زود “from (the) early morning”</td>
<td>صبح زود “(the) early morning”</td>
</tr>
<tr>
<td>سرانجام بعد از اتمام اعتصاب کارگران در روز پنجشنبه “finally after the end of the workers’ strike on Thursday”</td>
<td>روز پنجشنبه “Thursday”</td>
</tr>
<tr>
<td>طی دو سال گذشته “over the last two years”</td>
<td>دو سال گذشته “the last two years”</td>
</tr>
</tbody>
</table>

16 But see Section 3.4 on “Modified Temporal Expressions” for a discussion of prepositions that act as modifiers.
Note that certain prepositions in Persian are in fact compound elements, generally consisting of a noun and a preposition. Examples include "before", "after", "in/during". These compound elements should be treated as prepositions and should not be included in the annotated expression. However, as discussed in the last chapter, certain compound elements consisting of an adjective and a preposition that can be used to modify a temporal expression, such as "less than" and "more than", should be included within the extent of the annotation. Similarly, modifier expressions like "close to" and (when used to mean "up to") function more as modifiers of temporal expressions, and although they are prepositions, they should be included in the extent of the tag.

4.1.2 Morphological criteria

A number of elements such as pronouns, indefinite markers, and object markers, may be written as separate entities in Persian orthography or they can appear as affixes joined to the temporal expression. This section discusses which ones of these morphemes should be included within the extent of the annotation and which ones should be excluded.

The following are examples illustrating various elements such as pronouns, indefinite articles, relativizers, and definite articles, which should be included within the extent of the annotation if they are modifying a temporal expression. These elements are included in the annotation whether they appear as affixes attached to the previous element (e.g., the pronoun in the first example below) or as separate entities (as in the second example with ما “our”).

**Pronouns**

امروز روز تولدم بود.

*Today was my birthday.*

اینده یا اعتیاد و پی کاریست و یا در بهترین شکل مهاجرت از ایران!

*Our future is addiction and unemployment or in the best case scenario emigration from Iran!*

**Indefinite article or relativizer**

The indefinite article has the forms ی / ای and usually means “a/an” as in کتابی “a book”, or “some” if it is used with a plural noun as in کتابهایی “some books”. The relativizer has the same form as the indefinite article but is used on a noun or adjective when the latter is followed by که “that” which introduces a relative clause as in کتابی که من خریدم ...

17 The form ی / ای is also used to form adjectives as in مانندسالهای قرن بیست و پنجمی “us, twenty-first century humans”. In here, the ی affix is not an indefinite article meaning “a twenty first century” but rather turns the whole thing into
On such a day in 1285 AP Parvin Etessami the famous Iranian poet was born. The young poets have proclaimed their personal experiences with signs of a tumultuous century upon this decade’s poetry scene.

**Definite article**

The definite article is used only in the colloquial language as in "the book" or "that book":

من این تابستانی که شیباب اون زمستانه هست رو نمی خوام.

I don’t want this summer that is similar to that winter.

**4.1.2.1 Pronouns**

As already mentioned, certain elements such as possessive pronouns, if they are part of the temporal expression, should be tagged within the annotation. For instance, the pronoun "our" is included within the annotation of the temporal expression "our last week" as in its English counterpart. It is, however, not included in the annotation of "our last week’s meeting = our meeting of last week" because in this case, the pronoun is actually modifying the syntactic head of the full noun phrase, namely "meeting" (i.e., "our meeting" not "our week"). The two distinct annotations are illustrated below:

هفته گذشته مان

our last week = this last week of ours

< TIME2 VAL="2007-W38">هفته گذشته مان </TIME2>

جلسه هفته گذشته مان

our last week’s meeting = our meeting of last week

< TIME2 VAL="2007-W38">جلسه هفته گذشته مان </TIME2>

Note that at times the annotator may need to exclude an attached affix from the extent of the annotation. In certain cases, the meaning obtained may not be as easy to distinguish since the possessor pronoun is always placed at the end of the noun phrase in Persian. For instance, the

an adjective. As expected, it needs to be included within the annotation of the expression "twenty-first century".
sentence [Lit. poets of this century of our] is ambiguous based on whether it is referring to “our century” or “our poets”. The annotator should use the context and the meaning of the sentence in each case to determine the correct annotation. The two possible annotations are shown below:

The poets of this century of ours / The poets of our century [refers to “our century”]
< TIME2 VAL="2007-W38">شاعران این قرنمان</TIME2>

Our poets of this century [refers to “our poets”]
< TIME2 VAL="2007-W38">شاعران این قرنمان</TIME2>

4.1.2.2 Copula verb
The copula verb is the verb ‘to be’ in the present tense. The copula verb is never included within the annotation of the expression as shown in the following examples where the temporal expression is underlined and the copula is shown in bold:

امشب شش سال و چند ساعت است که او رفته است.
Tonight it’s six years and a few hours that he is gone.

گفت راستی نمی دونم اونا مال کودوم دهه اند؟ گفتند دویست نمیدونم شاید هفتاد یا هشتاد؟
She said hey don’t you know which decade they are from? I said not really maybe seventies or eighties.

The copula verb may appear attached to the preceding temporal element, but it should still be excluded from the extent of the annotation. Hence, in the first example below the suffix ـٗ "is" and in the second sentence the suffix ـُ "am" are not included in the annotation.

کمی بیش از دویست سال که با همسرم آزاده در شهر وینیپگ، کانادا، زنده می کنم.
It’s a bit more than two years that my spouse Azadeh and I are living in the city of Winnipeg, Canada.

شرم دارم که انسان این قرنم!!
I am ashamed that I am a human from this century.

In the following example, the final adjective has both an indefinite article and a copula verb (دردناک + ی + ـه); the indefinite marker is included but the copula verb is excluded from the annotation as in English:

زمستان سرد در دنیاکه.
It is a cold painful winter.

4.1.2.3 Object and topic marker را
In Persian, specific objects are marked with را as in “I bought the book”. The same affix can appear on topicalized elements (the topic of the discussion) as in این كتاب را، من پاریس خریدم which in English would translate to “As for this book, I bought it in Paris”. The affix را should never be included within the extent of the annotation, whether it appears attached to or detached from the temporal expression. Note that را can also be written as رو or in colloquial writing, in which case it is often joined to the preceding word.
These kids are forced to sell gum all day long in the metro, bus and the street corner, in search of a piece of bread.

The scene of those rejected at the door of the embassy made those two weeks hell for me...

I hope that you are feeling well and that this year’s Sizdah Bedar was spent well.

The Homa company has cancelled today’s flights.

### 4.2 Syntactic Criteria

The full extent of the tagged expressions includes all **premodifiers** of the time expression, among which are determiners (یک، چند، زمان، زمانه)، numerals (یک، دو، سه، چهار، پنج، شش)، quantifiers (طبیعی، بهترین)، as well as all **postmodifiers** in the noun phrase which in Persian include possessive noun phrases, pronouns, and all nominal/adjectival postmodifiers. For example, all of the following represent the full extent of markable temporal expressions:

<table>
<thead>
<tr>
<th>Persian</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>روز بعد</td>
<td>the next day</td>
</tr>
<tr>
<td>صبح زود</td>
<td>early in the morning</td>
</tr>
<tr>
<td>این روز های سرد زمستانی</td>
<td>these cold winter days</td>
</tr>
<tr>
<td>اواخر تابستان همون سال</td>
<td>early summer of that year</td>
</tr>
<tr>
<td>بهار آینده</td>
<td>next spring</td>
</tr>
<tr>
<td>دهه‌های اخیر</td>
<td>recent decades</td>
</tr>
<tr>
<td>نصف شب سعی‌نشین‌ها</td>
<td>midnight on Tuesdays</td>
</tr>
<tr>
<td>آینده خود را</td>
<td>its own future</td>
</tr>
<tr>
<td>آینده کشورمون</td>
<td>our country’s future</td>
</tr>
<tr>
<td>سال های ابتداي انقلاب</td>
<td>the early years of the revolution</td>
</tr>
<tr>
<td>غمگین ترین روز جهان</td>
<td>the saddest day in the world</td>
</tr>
<tr>
<td>مقصد ترین روز مذهبی</td>
<td>the holiest Jewish religious day</td>
</tr>
<tr>
<td>فقط یک سال پیش</td>
<td>just a year ago</td>
</tr>
<tr>
<td>تنها به ساعت</td>
<td>only one hour (long)</td>
</tr>
<tr>
<td>بیش از یک ماه</td>
<td>more than a month</td>
</tr>
<tr>
<td>دقیقا دو سال پیش</td>
<td>exactly two years ago</td>
</tr>
<tr>
<td>جند سال پیشتر</td>
<td>several years earlier</td>
</tr>
<tr>
<td>پنج ساله</td>
<td>five years old</td>
</tr>
<tr>
<td>سه هفته بعد</td>
<td>three weeks later</td>
</tr>
<tr>
<td>ساعتها پیش</td>
<td>hours earlier</td>
</tr>
<tr>
<td>دیشب در‌روز</td>
<td>late last night</td>
</tr>
<tr>
<td>کمی بیشتر ازدیس سال‌‌بعد</td>
<td>a bit more than ten years later</td>
</tr>
<tr>
<td>خیلی روزها</td>
<td>many days</td>
</tr>
<tr>
<td>یک ساعت و 10 دقیقه</td>
<td>an hour and 10 minutes</td>
</tr>
<tr>
<td>زودتر از موعد مقرر</td>
<td>before schedule</td>
</tr>
</tbody>
</table>

Note that what is referred to as adjective phrases in the English examples in the TIDES Standard for the Annotation of Temporal Expressions are often noun phrases in Persian as their structure is slightly different as in “a year ago” (یک سال پیش).
The full extent also includes all **postmodifiers** of the time expression, among which are adverbs, preposition phrases, and dependent clauses. For example, all of the following represent markable temporal expressions (in these examples, the heads are underlined and the postmodifiers are shown in italic):

<table>
<thead>
<tr>
<th>يک روز بعد از زلزله تهران</th>
<th>one day after the Tehran earthquake</th>
</tr>
</thead>
<tbody>
<tr>
<td>آینده جامعه ما</td>
<td>the future of our society</td>
</tr>
<tr>
<td>نزدیک به سه دهه حکومت ملایمان</td>
<td>close to three decades of the regime of the Mullahs</td>
</tr>
<tr>
<td>واپسین ماههای پیش از نابودی میهن فشنجم</td>
<td>the last months prior to the destruction of my beautiful country</td>
</tr>
<tr>
<td>همین هفت سال پیش، که میرعلیی را کشتند</td>
<td>just seven years ago when Mir-Alayi was murdered</td>
</tr>
<tr>
<td>یک روز تاریخی برای ایرانیان و فارسی زبانان</td>
<td>a historic day for Iranians and Persian speakers</td>
</tr>
</tbody>
</table>

In Persian, however the postmodifier may sometimes be separated from the temporal head since structural elements may move around more freely. As a general rule, the postmodifier should be included within the extent of the annotation of the temporal expression if it appears adjacent to the syntactic head, but not if it is separated by intervening elements. For instance, the sentence “It has been 20 years since the death of Gholamhossein Saedi” can be translated into Persian as in the first sentence below, where the temporal noun سال “year” and the postmodifier are adjacent and can be included within the same annotation:

بیست و دو سال از مرگ غلامحسین سعیدی گشته.

It has been twenty years since the death of Gholamhossein Saedi.

بیست و دو سال از مرگ غلامحسین سعیدی گشته. <TIME2 VAL="P20Y" ANCHOR_DIR="ENDING" ANCHOR_VAL="2007-10-22">

But the following structure where the verb intervenes between the temporal element and the postmodifier is quite common, especially in colloquial style. In this case, the postmodifier should be ignored in the annotation.

بيبست و دو سال گشته از مرگ غلامحسین سعیدی.

Twenty years have passed since the death of Gholamhossein Saedi.

بيبست و دو سال گشته از مرگ غلامحسین سعیدی. <TIME2 VAL="P20Y" ANCHOR_DIR="ENDING" ANCHOR_VAL="2007-10-22">

Non-adjacent temporal nouns and postmodifiers are especially common when the postmodifier is a relative clause as in the following examples where the intervening verb is shown in bold:

همین هفت سال پیش بود که میرعلیی را کشتند.

It was just seven years ago that Mir Allayi was murdered.
The wife of one of our relatives was talking about the first day that she had to take her daughter to one of these Islamic schools all by herself.

Note also that the inclusion of adjacent pre- and post modifiers in a temporal expression applies only to the modifiers of the time expression (i.e., the head of the phrase has to be itself a markable time expression). The time expression may be included within a larger phrase whose head is not part of the time expression; this head and the non-temporal modifiers of this head are not included within the scope of the TIME2 tag. For example, in the following, the head noun “message” and the postmodifier “of Ayatollah Khamenei” are not included in the TIME2 tag, and neither is the non-temporal premodifier of that string, namely the determiner “this.”

این پیام اخیر آیت الله خامنه‌ای به ملت بزرگ ایران است.
This is Ayatollah Khamenei’s recent message to the great people of Iran.

این پیام

In this next example, neither the noun “University students” after 2 “two-hour long” nor the preceding noun جلسهٔ “meeting (of)” is included within the scope of either TIME2 tag in Persian.

من امروز صبح به جلسهٔ ۲ ساعتی دانشجویان نتونستم برسم.
I was not able to make it to the two-hour long meeting of the University students this morning.

من امروز صبح به جلسهٔ دانشجویان نتونستم برسم.</TIME2>

جلسهٔ ۲ ساعتی

<TIME2 VAL="2007-10-22TMO">امروز صبح</TIME2>

من امروز صبح به جلسهٔ دانشجویان نتونستم برسم.
<TIME2 VAL="PT2H">۲ ساعته</TIME2>و<TIME2 VAL="2007-W46">من امروز صبح به جلسهٔ دانشجویان نتونستم برسم.

Similarly, the strings “picnic” and “in Black Strap” are not included within the scope of the TIME2 tag below:

عکسی از پیککینی هفته پیش در بلوک استرپ.
A photo from last week’s picnic in Black Strap.

عکسی از پیککینی هفته پیش در بلوک استرپ. <TIME2 VAL="2007-W46">هفته پیش</TIME2>

در بلوک استرپ.

Finally, in this last example, as the syntactic head is سن “age” and is not a temporal expression, the head noun and its modifier at least “at least” are not included in the extent of the annotation.

برای افتتاح حساب‌های سرمایه‌گذاری حداکثر سن ۱۸ سال تمام الزامی می‌باشد.
To establish investment accounts a minimum age of 18 years completed is required.

برای افتتاح حساب‌های سرمایه‌گذاری حداکثر سن الزامی می‌باشد. <TIME2 VAL="P18Y">۱۸ سال تمام</TIME2>
4.2.1 Appositives

An appositive to a temporal expression is considered a post-modifier and is thus included within the extent. If the appositive phrase itself contains a temporal expression, it is given a separate tag as well. In the following example, “the years of the chain murders” is an appositive that includes a temporal expression and it receives an embedded annotation.

The Seventies, the years of the chain murders, was one of the darkest points of modern history.

The verb-final word order in Persian may make it difficult to distinguish an appositive from a predicate element. An appositive is generally described as a noun phrase set beside another noun phrase to define, modify, or identify it. Appositives are therefore usually optional and can be removed from the sentence without affecting the sentence structure. In the two examples below, however, the underlined phrase is a predicate nominal and not an appositive. A predicate nominal describes the subject noun phrase, is linked to it by the verb ‘to be’, and is not optional. Hence, the subject noun phrase and the predicate nominal should receive separate TIME tags. Note that in English the predicate nominal appears following the verb.

Si Khordad Shust, Roz Aegaz Siae Terin Sal Hasting.

30 Khordad ’60, is the day of the beginning of one of the blackest years.

The ’60s, was the decade of big ideals.

4.2.2 Range Expressions

If the text of a temporal range expression has explicit begin and end points, then it is considered a range expression, and the points are tagged separately.

There exist few historical reports about the war in Kabul from 1992 through 1996.

There are no specific tags at this point for “workday” in TIME2.
Microsoft has announced that within the next three to four years this company will become one of the main players in the field of online advertising.

4.2.3 Conjoined Expressions

Phrases involving conjunction or disjunction of time expressions are handled similarly to ranges.

Many cases of conjoined expressions involve elision, but two tags should still be created:

البته در این شیوه حکومت منافقان آمریکایی ها بر منافق کوتاه و دیروز مدت افغان ها اولویت دارد.

Of course in this style of government the interests of Americans will take precedence over the short and long term interests of Afghans.
However, in cases like the following, create one tag and use the EQUAL_OR_MORE value in MOD.

If you have experienced long-lasting agitation and strong anxiety for a period of six months or more, you may have Generalized Anxiety Disorder.

In Persian, a conjunction may be represented by juxtaposing two numbers in order to indicate vagueness. For instance, “two three days” would mean “two or three days”. These should be tagged as two distinct conjoined expressions.

This section contains guidelines for determining whether a text segment that contains more than one hierarchically-related or offset-related temporal element should be assigned one tag or two tags. In the cases where two tags are expected, these guidelines also define the conditions for generating embedded versus non-embedded tags.

4.2.4 Embedded Expressions

An expression is treated as an indivisible syntactic unit if (a) there is no intervening element between temporal terms and (b) the two terms express values for units that are hierarchically related. For example:

جمعه شب

Friday evening (where the day unit is larger than and contains the part-of-day unit)

نوامبر 1943

November 1943 (where the month unit is smaller than and contained within the year unit)

Thus, each of the following expressions represents the full extent of a single TIME2 tag:
Friday evening
8:00 p.m. Friday
Tuesday the 18th of Azar
November 1943
the 12th of January of 1955 in the Gregorian calendar

Note that the familiar “day month year” format is considered a single TIME2 expression. A comma between TOD and Calendar Date expressions is ignored, and the whole is considered a single TIME2 expression in spite of the punctuation. Thus, the following are single TIME2 expressions (notice that the English versions of the date contain a comma, which is likewise ignored):

\[ \text{سمناطقٍ، 22 آبان 1386، 11:11:15} \]

\[ \text{ثُذّشّٖٖ، 5 نِٛٚٗٗ (5 November)} \]

Note however that if the expression includes two different calendars, they need to receive two distinct tags. In the following example, each underlined set of elements receives a separate annotation:

\[ \text{جسَمآردهم دسمير 2006 (23 آذر ماه)} \]

In Persian, the elements within a noun phrase are joined to each other using the ezaf suffix. English, on the other hand, often uses prepositions to link the noun phrase constituents. Hence, the constructions described in section 5.2.4.1 of the TIDES Standard for the Annotation of Temporal Expressions are expressed without prepositions and with the ezaf construction in Persian and therefore the same issues do not arise (also see Section 3.1 in this document). Hence, prepositions in Persian introduce syntactically embedded phrases and therefore typically require a separate TIME2 tag. There are, however, several instances in which you should ignore the preposition in Persian and create a single tag:

1. The preposition به “to” in time-of-day (TOD) expressions.
   
   \[ \text{ده دقیقه به سه} \]

2. The preposition از “from” in certain partitive constructions that are not formed with the ezaf.
   
   \[ \text{هر ساعتی از شبته روز} \]

3. The preposition بعد az “after” in TOD expressions in the night.
   
   \[ \text{سهان 30 بعد از نیمی شب} \]
   \[ \text{2:30 am [lit. 2:30 after midnight]} \]
4. Prepositions in rate expressions. These expressions can contain quantifiers, articles, or prepositions, so we treat them all equally. They are all SET expressions.

For instance, to say “(eight glasses of water) per day”, any of the following can be used:

- در روز
- در هر روز
- در شبانه روز
- در طول روز
- در کل روز

In certain time expressions, other elements besides a preposition may intervene such as a verbal participle. These elements are part of the time expression and should be included in the annotation:

پانزده دقیقه مانده به دو

4.2.4.2 When to Create Multiple Tags, with Embedding

There are two situations in which a TIME2 tag is embedded in the extent of another in Persian:

1. **Time-Anchored Expressions.** If a temporal expression includes an explicit anchor (i.e., if it explicitly expresses a temporal sequence), and if the anchoring phrase is itself a time expression, two tags are created, and the tag on the anchoring phrase is contained within the extent of the tag of the complete phrase. The value of the complete phrase is computed in relation to the value of the anchoring phrase.

You can return home two weeks after Tuesday’s operation.

دومهته بعد از عمل روز ساختنیه می‌توانید به خانه برگردد. چهار سال پیش در هفتم روز زمین لرزه‌ای در شهریم روی داد.

Four years ago today an earthquake took place in the city of Bam.

2. **Pre-Modifier and Head are Both Triggers.** When a collocation consists of two triggers that are not hierarchically related temporally, treat it as a compound noun, where the larger term is given one tag, and the pre-modifier(s) is/are given another.

His 20 Ordibehesht birthday

تولدش در روز 20 اردیبهشت
Possessive Constructions. As possessions are usually formed with the 'ezafe’ affix in Persian, the distinctions that arise in English cannot be found in this language. Hence, English can make a distinction between this year’s summer and the summer of this year and may annotate the two constructions differently. However, in Persian, both noun phrases will be represented simply as [summer-Ez this year] and will receive a single tag.

Also note that in Persian, as most time expressions include the specific time of day, the whole expression should receive a single extent. In these expressions, both the time and the time of day part are describing values for the same TOD unit and they are linked with the ezafe affix:

- ساعت دوازده شب (twelve o’clock at night)
- ساعت دوازده ظهر (twelve o’clock noon)
- صبح ساعت پنج (five o’clock in the morning)

4.2.4.3 When to Create Multiple Tags, without Embedding

In cases other than those described above, temporal phrases that appear in close proximity are tagged as independent phrases. In Persian, if the elements of a temporal expression are joined to each other with the ezafe, then in general the whole phrase should be treated as one unit even if the English translation may qualify for two distinct tags:

کنسرت ساعت هشت روز چهارشنبه خواهد بود.
Our concert will be at eight o’clock Wednesday

However, if a preposition or other parts-of-speech such as a verb intervene between the two temporal expressions, they should receive distinct tags:

کنسرت ساعت هشت در روز چهارشنبه خواهد بود.
Our concert will be at eight o’clock on Wednesday.

کنسرت ساعت سراسر روز چهارشنبه هشت.
Our concert is on Wednesday exactly at eight.